The MG Composite
A valid and reliable outcome measure for myasthenia gravis

Ted M. Burns, MD
Mark Conaway, PhD
Donald B. Sanders, MD
On behalf of the MG Composite and MG-QOL15 Study Group

ABSTRACT
Objective: To study the concurrent and construct validity and test-retest reliability in the practice setting of an outcome measure for myasthenia gravis (MG).

Methods: Eleven centers participated in the validation study of the Myasthenia Gravis Composite (MGC) scale. Patients with MG were evaluated at 2 consecutive visits. Concurrent and construct validities of the MGC were assessed by evaluating MGC scores in the context of other MG-specific outcome measures. We used numerous potential indicators of clinical improvement to assess the sensitivity and specificity of the MGC for detecting clinical improvement. Test-retest reliability was performed on patients at the University of Virginia.

Results: A total of 175 patients with MG were enrolled at 11 sites from July 1, 2008, to January 31, 2009. A total of 151 patients were seen in follow-up. Total MGC scores showed excellent concurrent validity with other MG-specific scales. Analyses of sensitivities and specificities of the MGC revealed that a 3-point improvement in total MGC score was optimal for signifying clinical improvement. A 3-point improvement in the MGC also appears to represent a meaningful improvement to most patients, as indicated by improved 15-item myasthenia gravis quality of life scale (MG-QOL15) scores. The psychometric properties were no better for an individualized subscore made up of the 2 functional domains that the patient identified as most important to treat. The test-retest reliability coefficient of the MGC was 98%, with a lower 95% confidence interval of 97%, indicating excellent test-retest reliability.

Conclusions: The Myasthenia Gravis Composite is a reliable and valid instrument for measuring clinical status of patients with myasthenia gravis in the practice setting and in clinical trials.

Neurology® 2010;74:1434–1440

GLOSSARY
AChR+ = acetylcholine receptor antibody positive; CI = confidence interval; MG = myasthenia gravis; MG-ADL = myasthenia gravis–activities of daily living scale; MG-MMT = myasthenia gravis–manual muscle test; MG-QOL15 = 15-item myasthenia gravis quality of life scale; MGC = Myasthenia Gravis Composite; MGFA = Myasthenia Gravis Foundation of America; MuSK+ = muscle specific tyrosine kinase positive; ROC = receiver operating characteristic.

Outcome measures for myasthenia gravis (MG) differ in ease of use, time to completion, what is being measured, how it is being measured, and who reports symptoms and impairments (i.e., patient-reported vs physician-reported).1-15 Most of these scales are unweighted but a few have weighted test item response options.3,7,9 The Myasthenia Gravis Composite (MGC) scale consists of test items that measure symptoms and signs of MG, with weighted response options (table 1).3 The individual test items of the MGC were selected from existing MG-specific scales based on their performance during 2 randomized, controlled clinical trials of mycophenolate mofetil for MG.3,16,17 The test items were selected so as to be meaningful to both the physician and the patient, frequently abnormal in patients with active disease, and responsive to clinical change.3 We present the results of validity and test-retest reliability analyses of the MGC in the patient care setting.

METHODS Validity testing of the MGC was conducted at 11 neuromuscular centers (9 in the United States and 2 in Europe) during the routine care of subjects with MG, following approval by each center’s ethical standards committee on human experimen-
tion. The investigators at each center have special expertise in MG management. Patients over 17 years of age with a clinical diagnosis of MG were recruited for the study if the treating physician predicted that they would be seen back for follow-up within 6 months. Written informed consent was obtained for each patient enrolled. The diagnosis of MG was made by the MG specialists at each center, and was almost always confirmed by serologic and electrophysiologic testing.

Prior to the study, the neck flexion test item of the original MGC was modified to include neck extension testing (worse score recorded). Also, physicians were asked to record the duration until development of ptosis or dysconjugate gaze to see if the time required for these 2 test items could be shortened. After learning that <1% of ocular score times were between 46 and 60 seconds, the ocular test items were modified so that 45 seconds was the maximum duration of testing.

At the initial visit, the following data were recorded: demographic information, disease characteristics (e.g., duration, antibody status, Myasthenia Gravis Foundation of America [MGFA] class), MG, and the Choose 2 component of the MGC. The Choose 2 form was administered by the physician, who asked the patient which 2 of 10 listed MG symptoms were considered by the patient to be most important to treat. A Choose 2 subscore was derived from the corresponding MGC scores of the patient’s 2 chosen items. Investigators at each site chose to also include either the MG-manual muscle test (MG-MMT) or MG–activities of daily living (MG-ADL) as part of their patient assessment at each visit. A gestalt impression of change was determined by the physician at the second visit, based on their overall assessment of the patient without considering change in any of these MG scale scores. The 15-item myasthenia gravis quality of life scale (MG-QOL15) was completed by the patient at each study visit.

Table 1  The Myasthenia Gravis Composite scale

<table>
<thead>
<tr>
<th>Test item (physician examination)</th>
<th>Score 1</th>
<th>Score 2</th>
<th>Score 3</th>
<th>Score 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pupils, upward gaze (physician examination)</td>
<td>&gt;45 seconds = 0</td>
<td>11–45 seconds = 1</td>
<td>1–10 seconds = 2</td>
<td>Immediate = 3</td>
</tr>
<tr>
<td>Double vision on lateral gaze, left or right (physician examination)</td>
<td>&gt;45 seconds = 0</td>
<td>11–45 seconds = 1</td>
<td>1–10 seconds = 3</td>
<td>Immediate = 4</td>
</tr>
<tr>
<td>Eye closure (physician examination)</td>
<td>Normal = 0</td>
<td>Mild weakness (can be forced open with effort) = 1</td>
<td>Moderate weakness (can be forced open easily) = 1</td>
<td>Severe weakness (unable to keep eyes closed) = 2</td>
</tr>
<tr>
<td>Talking (patient history)</td>
<td>Normal = 0</td>
<td>Intermittent slurring or nasal speech = 2</td>
<td>Constant slurring or nasal but can be understood = 4</td>
<td>Difficult to understand speech = 6</td>
</tr>
<tr>
<td>Chewing (patient history)</td>
<td>Normal = 0</td>
<td>Fatigue with solid food = 2</td>
<td>Fatigue with soft food = 4</td>
<td>Gastric tube = 6</td>
</tr>
<tr>
<td>Swallowing (patient history)</td>
<td>Normal = 0</td>
<td>Rare episode of choking or trouble swallowing = 2</td>
<td>Frequent trouble swallowing, e.g., necessitating changes in diet = 5</td>
<td>Gastric tube = 6</td>
</tr>
<tr>
<td>Breathing (thought to be caused by MG)</td>
<td>Normal = 0</td>
<td>Shortness of breath with exertion = 2</td>
<td>Shortness of breath at rest = 4</td>
<td>Ventilator dependence = 9</td>
</tr>
<tr>
<td>Neck flexion or extension (weakest) (physician examination)</td>
<td>Normal = 0</td>
<td>Mild weakness = 1</td>
<td>Moderate weakness [i.e., −50% weak, ±15%] = 3a</td>
<td>Severe weakness = 4</td>
</tr>
<tr>
<td>Shoulder abduction (physician examination)</td>
<td>Normal = 0</td>
<td>Mild weakness = 2</td>
<td>Moderate weakness [i.e., −50% weak, ±15%] = 4a</td>
<td>Severe weakness = 5</td>
</tr>
<tr>
<td>Hip flexion (physician examination)</td>
<td>Normal = 0</td>
<td>Mild weakness = 2</td>
<td>Moderate weakness [i.e., −50% weak, ±15%] = 4a</td>
<td>Severe weakness = 5</td>
</tr>
</tbody>
</table>

*Moderate weakness for neck and limb items should be construed as weakness that equals roughly 50% ±15% of expected normal strength. Any weakness milder than that would be mild and any weakness more severe than that would be classified as severe.

RESULTS Test-retest reliability. A total of 38 patients with MG were evaluated independently on the same day by 2 neurologists (T.M.B. and another), who were blinded to the results of the other evaluation. The order of evaluators varied depending on the situation; often the other evaluator (e.g., clinical neurophysiology fellow) performed the first evaluation followed by T.M.B., but other times T.M.B. evaluated first followed by a resident or fellow. The mean MGC score at the first assessment was 4.9 (SD = 7.1; range = 0–35). The MGC score never differed by more than 4 points between evaluators; it

...
A total of 175 patients with MG were enrolled at 11 sites from July 1, 2008, to January 31, 2009. A total of 151 patients were seen in follow-up (August 18, 2008, to August 1, 2009). The mean age at visit 1 was 58 years, and the mean disease duration was 7 years. A total of 53% were men and 47% were women. The mean time between visit 1 and visit 2 was 4.7 months (0.13–10.1 months). The 24 patients not seen at visit 2 did not differ from the other 151 patients in age ($p = 0.85$), MGC score at visit 1 ($p = 0.73$), MG-QOL15 score at visit 1 ($p = 0.21$), or gender ($p = 0.65$).

Breakdown by current MGFA Class at visit 1 is as follows: 22% Class I (ocular), 41% Class II (mild, generalized), 19% Class III (moderate, generalized), 3% Class IV (severe, generalized). A total of 15% of patients were in remission (Class 0) at visit 1. None were intubated and ventilated (Class V) at visit 1, but 14.3% had experienced at least 1 prior episode of myasthenic crisis. At visit 2, 24% were Class I, 42% Class II, 13% Class III, 0% Class IV, and Class V, and 20% were in remission. A total of 78% of patients were acetylcholine receptor antibody positive (AChR+) and 7% were muscle specific tyrosine kinase positive (MuSK+). A total of 10% of patients had negative serology for at least AChR (often both AChR and MuSK). Serologic status was unknown for 5%. Of the 175 subjects, 142 (82%) were follow-up outpatients, 23 (13%) were newly diagnosed, and 8 (5%) were hospitalized for MG at the time of the evaluation. A total of 110/175 (63%) patients were taking pyridostigmine at the time of their first visit.

The mean, median, SD, and range of MG scale scores for both visits are listed in table 2. The total MGC scores showed excellent concurrent validity with other MG-specific scales. At visit 1, the total MGC score had a correlation of 0.68 (95% CI 0.59–0.75) with the MG-QOL15 total score, 0.85 (95% CI 0.77–0.90) with the MG-ADL total score, and 0.80 (95% CI 0.72–0.86) with the MG-MMT total score. Nearly identical correlations were observed at visit 2.

Responsiveness and longitudinal construct validity of the MGC were assessed by using different gold standards of clinical change (table 3). The figure 1 demonstrates the ROC curve for the MGC using the physician impression of improvement plus improvement in MG-QOL15 score (visit 2–visit 1) as the criterion for improvement. Analyses of sensitivities and specificities of the MGC revealed that a 3-point improvement in MGC score was optimal for signifying clinical improvement. Furthermore, a ≥3-point improvement in MGC score was meaningful to most patients. For instance, of the 49 patients with a ≥3-point improvement in the MGC, 22% had negative serology for at least AChR (often both AChR and MuSK). Serologic status was unknown for 5%. Of the 175 subjects, 142 (82%) were follow-up outpatients, 23 (13%) were newly diagnosed, and 8 (5%) were hospitalized for MG at the time of the evaluation. A total of 110/175 (63%) patients were taking pyridostigmine at the time of their first visit.

### Table 2: Myasthenia Gravis Composite scale scores at visits 1 and 2

<table>
<thead>
<tr>
<th>Visit 1, all patients</th>
<th>No.</th>
<th>Mean</th>
<th>SD</th>
<th>Med</th>
<th>Min–max</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGC total</td>
<td>175</td>
<td>7.6</td>
<td>6.7</td>
<td>6</td>
<td>0–28</td>
</tr>
<tr>
<td>MG-ADL total</td>
<td>87</td>
<td>4.9</td>
<td>3.5</td>
<td>5</td>
<td>0–13</td>
</tr>
<tr>
<td>MG-QOL15 total</td>
<td>175</td>
<td>18.7</td>
<td>14.8</td>
<td>17</td>
<td>0–52</td>
</tr>
<tr>
<td>MG-MMT total</td>
<td>107</td>
<td>6.3</td>
<td>6.8</td>
<td>4</td>
<td>0–41</td>
</tr>
<tr>
<td>Choose 2 total</td>
<td>175</td>
<td>3.0</td>
<td>2.6</td>
<td>3</td>
<td>0–10</td>
</tr>
<tr>
<td>Visit 1, newly diagnosed patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGC total</td>
<td>23</td>
<td>9.7</td>
<td>6.2</td>
<td>8</td>
<td>0–22</td>
</tr>
<tr>
<td>MG-ADL total</td>
<td>8</td>
<td>7.3</td>
<td>3.4</td>
<td>6</td>
<td>4–13</td>
</tr>
<tr>
<td>MG-QOL15 total</td>
<td>23</td>
<td>26.6</td>
<td>14.0</td>
<td>28</td>
<td>0–52</td>
</tr>
<tr>
<td>MG-MMT total</td>
<td>15</td>
<td>8.5</td>
<td>5.2</td>
<td>8</td>
<td>2–23</td>
</tr>
<tr>
<td>Choose 2 total</td>
<td>23</td>
<td>4.3</td>
<td>2.7</td>
<td>5</td>
<td>0–7</td>
</tr>
<tr>
<td>Visit 1, MGFA class ≥2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGC total</td>
<td>110</td>
<td>10.5</td>
<td>6.6</td>
<td>10</td>
<td>0–28</td>
</tr>
<tr>
<td>MG-ADL total</td>
<td>60</td>
<td>6.2</td>
<td>3.3</td>
<td>6</td>
<td>0–13</td>
</tr>
<tr>
<td>MG-QOL15 total</td>
<td>110</td>
<td>24.3</td>
<td>14.2</td>
<td>23</td>
<td>1–52</td>
</tr>
<tr>
<td>MG-MMT total</td>
<td>63</td>
<td>8.9</td>
<td>7.4</td>
<td>7</td>
<td>0–41</td>
</tr>
<tr>
<td>Choose 2 total</td>
<td>110</td>
<td>3.7</td>
<td>2.4</td>
<td>4</td>
<td>0–10</td>
</tr>
</tbody>
</table>

Abbreviations: MG-ADL = myasthenia gravis-activities of daily living scale; MG-MMT = myasthenia gravis-manual muscle test; MG-QOL15 = 15-item myasthenia gravis quality of life scale; MGC = Myasthenia Gravis Composite; MGFA = Myasthenia Gravis Foundation of America.
Using the physician impression of improvement plus improvement in the MG-QOL15 score as the gold standard for clinical improvement, the effect on sensitivity and specificity of different choices for cut-offs for the MGC, MG-ADL, and MG-MMT scores is shown in table 4. A 1- or 2-point drop in the MGC also provides high sensitivity with small sacrifices in specificity. The sensitivities and specificities for the MGC compare favorably with the MG-ADL and MG-MMT (table 4).

In the subgroup of patients with initial MGC scores of 5 or more, the sensitivity of a 3-point change in MGC was 94.6% (n = 42) and the specificity was 80.0% (n = 45). In the group with initial MGC score of 7 or more, the sensitivity was 100% and the specificity was 81%.

Analysis of the Choose 2 component of the MGC. Eyelid droop (33%), double vision (36%), and trouble swallowing (36%) were the 3 most common domains selected by the patient as being most important during the Choose 2 component of the MGC. At each visit, the correlations for the MG-MMT, MG-ADL, and MG-QOL15 to the Choose 2 subscale score were lower than to the total MGC score. For all the gold standards of change listed in table 3, the Choose 2 component showed either lower sensitivity or lower specificity than the MGC total (data not shown). These results do not support the use of the Choose 2 component of the MGC over the total MGC score.

DISCUSSION The notable attributes of the MGC include the following. 1) The test items were rigorously selected through a process that assessed item performance during 2 randomized, controlled trials.
The sample enrolled in this study is likely representative of the overall population of patients with MG because it included consecutive patients seen for routine care at 11 centers. The 9 centers in the United States were located in 9 different states (4 East Coast, 4 Midwest, and 1 West Coast). The 2 centers in Europe were in Glasgow and Milan. Our sample population is similar in many characteristics to other cohorts of patients with MG and covered the spectrum of MG severity. For example, the serologic status, percentage of patients with previous crisis, and proportion of patients by MGFA Class were similar to reported cohorts said to be representative of MG.21,22 Also, the mean MG-ADL score at visit 1 in our cohort was similar to what it was for patients enrolled in the Aspreva MG trial.23 However, overall severity in our cohort was probably less than that of patients with MG entering most clinical trials because of exclusion from those trials of patients with milder disease (e.g., ocular MG [MGFA Class I] or patients in remission). For example, the mean MG-MMT (6.3) and MG-ADL (4.9) scores in our patients were lower than in the recent Muscle Study Group trial of mycophenolate mofetil for generalized MG.

The MGC can be used in everyday practice and in clinical trials. For clinical trials, we propose that a 3-point fall in MGC score be the cutpoint to indicate clinical improvement. The 3-point threshold demonstrated acceptable test-retest reliability and offered the best trade-off in terms of sensitivity and specificity when compared with the various criteria chosen to signify clinical improvement. A 3-point improvement in MGC also appears to be meaningful to most patients, as evidenced by a mean improvement in MG-QOL15 score of >12 points for those patients who had a ≥3-point improvement in the MGC.

The MGC differs from most scales in that it is a hybrid of physician-reported and patient-reported test items. This combination, however, is especially appropriate for an MG outcome measure for 3 major reasons. 1) MG manifests with fluctuating and fatigable symptoms and signs, which consequently limits the utility of measuring disease status solely based on a “snapshot” examination. 2) Manifestations of MG are evident not only to the physician but also to the patient (and others), and thus patient-reported measures of symptoms must be considered in assessing disease severity. 3) Dysfunction in chewing and swallowing are almost certainly more evident to the patient than to the physician. It is not surprising that certain patient-reported test items perform better than their physician-reported examination counterparts and thus justify inclusion in the MGC.

A limitation of our study is the lack of a single gold standard for clinical change. On the other hand, if there had been a gold standard then there would have been less reason to develop the MGC in the first place. To overcome this limitation of no gold standard, we looked at the performance of the MGC using several different criteria of change, each of which has its own limitations. For example, one problem with using MG-MMT or MG-ADL to assess MGC responsiveness is that the MGC contains 4 test items from each the MG-MMT and the MG-ADL. However, it was not reasonable to remove these shared items for the validity analysis because each item represents a functional domain that often acts independently of the other domains. Thus, assessing correlations with scales stripped of shared test items would offer little insight into the performance of the MGC. Limitations in using the physician impression of change as a comparator included physician uncertainty, limited physician memory of patient status at...
visit 1, and the fact that in some cases different physicians saw the patient at the 2 visits. We settled on the physician impression of improvement plus improvement in the patient-derived MG-QOL15 score as our gold standard for clinical improvement, and placed the most weight on this comparator to assess the longitudinal validity of the MGC and to determine the optimal cutpoint for signifying clinical improvement.

COINVESTIGATORS
MG Composite and MG-QOL15 Study Group: Principal Investigator: Ted M. Burns, MD (University of Virginia); Investigators: Mark Conaway, PhD, Guillermo Solorzano, MD (University of Virginia); Maria E. Farrugia, DPhil, MD (Institute of Neurological Sciences, Glasgow, UK); Donald B. Sanders, MD, Janice M. Massey, MD, Vern C. Juel, MD, Lisa D. Hobson-Webb, MD, Bernadette Tucker-Lipson, RN, MSN (Duke University); Carlo Antozzi, MD, Renato Mantegazza, MD (Neurological Institute Foundation “Carlo Besta,” Milan, Italy); David Lacomis, MD (University of Pittsburgh); Elliot Dimberg (Mayo Clinic, Florida); Srijanik Muppudi, MD, Gil I. Wolfe, MD (University of Texas, Southwestern); Mazen M. Dimachkie, MD, Richard J. Barohn, MD, Mamatha Pasnoor, MD, April L. McVey, MD, Laura Herbelin, BS (University of Kansas); Taheen Mozaffar, MD, Vish Q. Dang, MD, Sandhya Rao, MD (University of California, Irvine); Robert Pascuzzi, MD, Riley Snook, MD (Indiana University); Anthony A. Amato, MD (Brigham and Women’s Hospital and Harvard Medical School).

DISCLOSURE
Dr. Burns serves as Section Editor of the Neurology® Podcast and receives research support from Alexion Pharmaceuticals, Inc., Knopp Neurosciences Inc., and the Myasthenia Gravis Foundation of America. Dr. Sanders has served/serves as a consultant for Accordant Health Service, Bayhill Therapeutics, BioMarin Pharmaceutical Inc., Anthera Pharmaceuticals, Jacobus Pharmaceutical Co. Inc., Baxter International Inc., and GalAsthm/Kline, and serves on speakers’ bureaus for Athena Diagnostics, Inc. and Discovery Institute of Medical Education/Talecris Biotherapeutics. Dr. Conaway serves on the editorial board of Developmental Neurobiology and Child Neurology and has received research support from the Myasthenia Gravis Foundation of America. Dr. Farrugia reports no disclosures. Dr. Solorzano’s wife has received speaker honoraria from Pfizer Inc. and she has received research support from Novo Nordisk, the NIH (2U54HD828934-16A1 revised [Co-I]), and the University of Virginia Children’s Hospital. Dr. Massey serves on the editorial board of Seminars in Neurology; receives an educational grant from Allergan, Inc. for an annual workshop; receives research support from the NIH (ADD GRANT NUMBER [Site PI]); and serves on Board of Directors of the American Board of Psychiatry & Neurology (ABPN)—Neurology. Dr. Juel receives research support from Alexion Pharmaceuticals, Inc. Dr. Hobson-Webb has received honoraria for speaking or educational activities not funded by commercial entities and receives research support from Genzyme Corporation and the American Association of Neuromuscular and Electromyographic Medicine. Dr. Tucker-Lipson has received research support from Alexion Pharmaceuticals, Inc. and the NIH/UEAB. Dr. Antozzi serves on the editorial board of Neurological Sciences. Dr. Mantegazza serves on a scientific advisory board for GlaxoSmithKline; has received funding for travel from Sanofi-Aventis; has received a speaker honorarium from MEDA Pharmaceuticals Inc.; and receives research support from the FDA for the NIH, as principal investigator for 2 studies: 1R01FD003538 Phase II therapeutic trial of mexiletine in patients with non-dystrophic myotonia and 1R01FD003454 Phase II trial of methylxothiazole in myasthenia gravis. Dr. McVey and Dr. Pasnoor report no disclosures. L. Herbelin has serves as a consultant for Alexion Pharmaceuticals, Inc. and has received research support from support from the FDA, Orphan Products Division (1R01FD003538 [Project manager]). Dr. Mozaffar serves on the Medical and Scientific Advisory Board for the Myositis Association; has received funding for travel from and on speakers’ bureaus for Genzyme Corporation and Talecris Biotherapeutics; has served as a consultant for Avanir Pharmaceuticals and Baxter International Inc.; serves as Medical Director for Crescent Health Care; and receives research support from the NIH (RO1SR01NS049203 [Co-I]). Dr. Dang, Dr. Rao, and Dr. Snook serve/serves no disclosures. Dr. Amato has served on the Medical Advisory Board for MedImmune; serves as an Associate Editor for Muscle & Nerve and on the editorial board of Neurology®; receives royalties from the publication of Neuromuscular Disease (McGraw-Hill, 2008); and receives research support from MedImmune and the NIH (R01 NS04963-9-01A2 [PI] and U54 RR019482 [Co-I]).

Received November 17, 2009. Accepted in final form February 3, 2010.

REFERENCES
9. Mantegazza R, Antozzi C, Pelucchetti D, Ghirlanda Z, Cornelio F. Azathioprine as a single drug or in combina-


The MG Composite: A valid and reliable outcome measure for myasthenia gravis
Ted M. Burns, Mark Conaway and Donald B. Sanders

Neurology 2010;74;1434-1440
DOI 10.1212/WNL.0b013e3181dc1b1e

This information is current as of May 3, 2010