Received: 28 October 2022 | Revised: 4 April 2023 | Accepted: 11 April 2023

DOI: 10.1002/mus.27837

CLINICAL RESEARCH ARTICLE

MUSCLE&NERVE WILEY

Concordance between patient- and physician-reported Myasthenia Gravis Activities of Daily Living (MG-ADL) scores

Sarah Dewilde PhD[1] | M. F. Janssen PhD[2] | Nafthali H. Tollenaar BBA[1] | Fiammetta Vanoli MD[3,4] | Rita Frangiamore MD[3] | Glenn Phillips PhD[5] | Sandra Paci PhD[5] | Renato Mantegazza MD[3] | Andreas Meisel MD[6,7,8,9] | Frauke Stascheit MD[6,7]

[1]Services in Health Economics, Brussels, Belgium

[2]Department of Psychiatry, Erasmus Medical Center, Rotterdam, The Netherlands

[3]Fondazione IRCCS Istituto Neurologico Carlo Besta Milan, Italy

[4]Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy

[5]Health Economics & Outcomes Research, argenx BV, Boston, Massachusetts, USA

[6]Department of Neurology, Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[7]NeuroCure Clinical Research Center, Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin, Germany

[8]Center for Stroke Research Berlin, Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin, Germany

[9]Integrated Myasthenia Gravis Center, Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health, Berlin, Germany

Correspondence
Sarah Dewilde, Services in Health Economics, Rue Jean Gerard Eggerickx 36, 1150 Brussels, Belgium.
Email: sd@she-consulting.be

Abstract

Introduction/Aims: Myasthenia gravis (MG) is a neuromuscular disease characterized by abnormal skeletal muscle fatiguability. The MG Activities of Daily Living (MG-ADL) scale assesses eight symptoms and is often used as primary endpoint in MG clinical trials where it is completed by neurologists. However, in observational studies, patients frequently complete the MG-ADL scale independently of their neurologist. In this study we aimed to assess the concordance between self- and physician-reported MG-ADL scores.

Methods: An international observational study was conducted among adult patients with MG scheduled for a routine visit or who entered the hospital via emergency services. Consenting patients and physicians completed the MG-ADL. Concordance between assessments was calculated using Gwet's agreement coefficient (Gwet's AC) for the MG-ADL individual items and the intraclass correlation coefficient (ICC) for the MG-ADL total score.

Results: Data were collected from 137 patients (63% female; mean age, 57.7 years). Physicians assessed the patient's symptoms as slightly more severe (8.1 vs 7.5 MG-ADL total score, respectively), corresponding to a difference of 0.6 on a range from 0 to 24. The ICC for the MG-ADL total score between the patient and the physician assessment was 0.94 (95% confidence interval, 0.89 to 0.95), showing excellent concordance. Gwet's AC showed substantial to almost perfect agreement for all items, except eyelid droop, for which the agreement was moderate.

Discussion: Our results demonstrate that patients and neurologists have a concordant assessment of the patient's MG symptoms when using the MG-ADL scale.

Abbreviations: AChR, acetylcholine receptor antibody status; CI, confidence interval; ICC, intraclass correlation coefficient; IVIg, intravenous immunoglobulin; MG, myasthenia gravis; MG-ADL, Myasthenia Gravis Activities of Daily Living; MGFA, Myasthenia Gravis Foundation of America; MID, minimal important difference; Q1, Q3, first quartile, third quartile; RCT, randomized clinical trial; SD, standard deviation.

Renato Mantegazza, Andreas Meisel, and Frauke Stascheit contributed equally to this study.

Funding information
argenx BV

This evidence supports patient self-administration of the MG-ADL in clinical practice and research.

KEYWORDS
concordance, MG-ADL, myasthenia gravis, observational study, proxy assessment

1 | INTRODUCTION

Myasthenia gravis (MG) is a chronic autoimmune neuromuscular disease characterized by extensive muscle fatigue, and is most commonly caused by antibodies against the muscle acetylcholine receptor (AChR).[1] In randomized clinical trials (RCTs) focused on MG treatment, the Myasthenia Gravis Activities of Daily Living (MG-ADL) total score[2] is often used as the primary endpoint to assess symptom severity (Table S1 in Data S1).[3–5] The MG-ADL can be completed within 2 or 3 minutes and requires no equipment or formal training. In clinical studies, patients are assessed by a neurologist who scores the MG-ADL scale while examining and interviewing them. In observational longitudinal studies, data on patients' health status are often collected at home from a wide range of patients treated with a broad range of therapies, and with different follow-up protocols. In these observational studies, patients describe their health-related quality of life and their utilization of medical resources and usually self-assess their MG symptoms using available patient-reported outcome measures, such as the MG-ADL. Both sources of data (RCTs and observational studies) are typically combined in pharmacoeconomic analyses. In addition, the MG-ADL could serve as a simple tool for telemedicine monitoring of patients. It is unclear how the variable use and administration of the MG-ADL in clinical care and research affects the score.

Therefore, the objective of this study was to assess the concordance between physician- and patient-reported assessments of the eight items from the MG-ADL scale and its total score. In addition, we assessed whether any differences that occurred were systematic and are linked to patient characteristics (e.g., age, sex, comorbidities) or disease severity (e.g., according to disease classification or antibody status).

2 | METHODS

2.1 | Study design and patient sample

A multicenter, cross-sectional, observational study was conducted among adult patients with MG. The study was carried out in two medical centers: the IRCCS Istituto Neurologico Carlo Besta, Milan, Italy, and Charité—Universitätsmedizin Berlin, Germany, where data were collected from March 2021 to December 2021. Ethics approval was obtained from the local ethics committees. At each medical center, one single physician tested and assessed all patients.

Patients were either recruited during routine scheduled appointments with their neurologist or when they had an unscheduled emergency visit due to an exacerbation. The inclusion criteria were: (1) being previously diagnosed with MG according to the current national guidelines[6]; and (2) age ≥18 years. No other inclusion or exclusion criteria were applied. Written informed consent was obtained from all participants.

A suitable spread in severity across MG-ADL scores was considered important to ensure a robust self-assessment vs clinician assessment concordance. Therefore, we opted for stratified recruitment using the last known MG-ADL or the Myasthenia Gravis Foundation of America (MGFA) clinical classification at the time of study inclusion. The stratification into mild (MGFA I or II, or MG-ADL 1 to 6), moderate (MGFA III or MG-ADL 7 to 15), or severely affected patients (MGFA IV or V, or MG-ADL >15) ensured that patients from the whole severity continuum were included in the concordance assessment.

2.2 | Data collection

Disease-specific data, such as AChR antibody status, muscle-specific kinase (MuSK) antibody status, ocular vs generalized MG, comorbidities, MG crisis in the past year (yes/no), thymectomy (yes/no), current treatment, MGFA class, and proxy MG-ADL information, were provided by the physician. Patients provided demographics (age, sex, weight, and height) and a self-assessed MG-ADL score. They were also asked whether their health status had changed between the day of their neurologist visit and the day they completed the MG-ADL. This question was added to verify whether the patient and their physician were assessing the same health state. No further explanation was provided regarding the nature of this health change.

To rule out the possibility of an order effect on the assessment, the order of administration of the questionnaires by the patients and physicians was randomized. A 2-day interval was recommended in the study protocol between the completion of a patient's and physician's assessment to minimize the likelihood of patients and physicians being influenced by each other's responses, and to minimize the likelihood of significant changes in the patient's health status. However, a maximum of 6 days (less than 1 week) between the two assessments was still considered acceptable for inclusion in the analysis, as the aim was to test the interrater agreement under various circumstances.

2.3 | Statistical analysis

Based on a type 1 error of 0.05, a power of 0.80, and previous research[7] indicating an interrater agreement of over 0.94 between

patient and physicians in MG-ADL scoring, the minimum sample size needed per stratum for our analysis was 35, indicating a minimum sample of 105.[8] Allowing for some missing data, we therefore planned on recruiting a 150-person sample to estimate the concordance between patient- and physician-assessed MG-ADL. Initially, we aimed for equal distribution across the severity spectrum. However, due to the difficulty of recruiting high-severity patients and their low prevalence, this ratio was adjusted to 40/40/20 for mild, moderate, and severe MG, respectively.

The distributions of MG-ADL items were calculated for both self-report and physician measurements. For the MG-ADL total score, parametric statistics were used (mean and standard deviation [SD]). Gwet's agreement coefficient (Gwet's AC) was used to assess the concordance of the MG-ADL individual items.[9] Gwet's AC is a statistic that addresses the paradoxical behavior of Cohen's kappa in cases where a low kappa is observed with a high agreement.[10] This measure was considered to be slight (0 to 0.20), fair (0.21 to 0.40), moderate (0.41 to 0.60), substantial (0.61 to 0.80), or almost perfect (>0.80).[11] Intraclass correlation coefficients (ICCs) were used to assess the concordance of MG-ADL total scores from different methods of administration. We used two-way, random effects, and average measure ICCs. ICCs were considered either poor (<0.5), moderate (0.51 to 0.75), good (0.76 to 0.90), or excellent (>0.90).[12] Concordance was also calculated for subgroups defined by changes in health status (yes/no) and by country. Differences between patient- and physician-reported scores were calculated and reported for the total study population by age, sex, MGFA class at the time of study inclusion, thymectomy (yes/no), antibody status, and number of comorbidities. Ordinary least-squares regression was used for testing the significance of these differences. We also tested for the effect of order of administration (self-assessment first vs proxy assessment first). All analyses were conducted using Stata version 16.1.[13] $P < 0.05$ was considered statistically significant.

3 | RESULTS

3.1 | Demographics

A total of 146 patients diagnosed with MG at two centers were enrolled in this study. Eight patients failed to complete the questionnaire within 6 days of the physician's assessment (5.5%), and one had missing data on one or more of the eight items of the MG-ADL (0.7%). The majority of the nine removed patients were women and were younger than the mean age of the sample. Therefore, the study sample consisted of 137 patients (63% female) with a mean (SD) age of 57.7 (17.8) years (Table 1). All patients had comorbidities, with cardiovascular and respiratory diseases being the most frequently occurring (Table S2). The mean number of days between patient and physician assessment was 1.8 (SD, 1.0; median, 2).

3.2 | Distribution of patient- and physician-assessed MG-ADL scores

Physicians assessed the patient's health status slightly more severely (0.6-point higher total score; Table 2). Despite the standard deviation of 2.3, this difference was ≤2 points in 80% of cases. The difference in mean MG-ADL scores between physicians' and patients' assessments was larger in Germany than in Italy. The MG-ADL total scores by patients and physicians did not show a structural under- or overreporting at a specific item or severity level.

3.3 | Concordance between patient- and physician-assessed MG-ADL total scores

The overall ICC for the total MG-ADL score indicated excellent agreement (Table 3). Interrater agreement of the MG-ADL individual items also indicated moderate to substantial agreement, as shown by Gwet's AC values ranging between 0.46 (eyelid droop) and 0.77 (chewing).

Figure 1A shows a bubble plot of the observed MG-ADL scored by patients vs physicians, which shows only a few outliers. Similarly, Figure 1B illustrates that most differences between the patient- and physician-assessed MG-ADL total score lie around zero, indicating excellent agreement, but the differences became larger and occurred more often in patients with higher MGFA class, that is, those with more severe disease.

Subtraction of the MG-ADL total scores (physician − patient) revealed that the majority of values per variable and item were just above zero (Table 4), indicating that, on average, the physicians judged the patient's illness to be similarly or somewhat more severe. Physicians tended to give higher (more severe) total scores to patients with increasingly severe MGFA classes ($P < 0.0001$), to patients who had AChR antibodies ($P < 0.05$), and to patients with an increasing number of comorbidities ($P < 0.01$). No differences in the MG-ADL total score were found in the subgroups defined by age, sex, or thymectomy. An exception was found in patients with severe disease (MGFA class IVa and IVb), whose MG-ADL total score was 2.00 to 4.43 points higher, on average, when assessed by a physician compared with the self-assessment. In these two MGFA classes, five of eight items contributed to this difference in total score. Some individual item-score differences were also significant between subgroups defined by the number of comorbidities. There was no significant effect on the MG-ADL total score or any individual MG-ADL items by order or randomization (whether the patient or physician filled in the MG-ADL first; results not shown).

Of the 137 patients, 18 (13.1%; 9 in Italy and 9 in Germany) indicated that they had experienced a change in their health status between their own assessment and the neurologist visit. As expected, a significantly lower ICC was found among patients who experienced a change in health between the proxy and the self-assessment compared with those who did not, and Gwet's AC was also lower for all items (Table 5). The ICCs were almost identical for the medical centers

TABLE 1 Patients' characteristics.

		Germany	Italy	Total
Number (%)		**95**	**42**	**137**
Age (years)	18–34	13%	14%	13%
	35–54	26%	33%	28%
	55+	61%	52%	58%
Sex	Female	57%	74%	63%
Years since initial diagnosis	<1 year	7%	12%	8%
	1–5 years	39%	21%	34%
	>5 years	54%	67%	58%
AChR	Positive	76%	74%	76%
MuSK	Positive	5%	18%	9%
MGFA class	I	14%	2%	10%
	II	36%	29%	34%
	III	32%	50%	37%
	IV	17%	12%	15%
	V	1%	0%	1%
	Unknown	1%	7%	3%
MG crisis over the last year	Yes	24%	26%	25%
Health change over last 2 days	Yes	9%	21%	13%
Thymectomy	Yes	46%	45%	46%
Current treatment[a]	Anticholinesterase medication	89%	71%	84%
	Corticosteroids	57%	81%	64%
	Azathioprine	29%	17%	26%
	Mycophenolate	18%	17%	18%
	Methotrexate	6%	2%	5%
	Eculizumab	15%	2%	11%
	Rituximab	4%	0%	3%
	Plasma exchange	8%	14%	10%
	IVIg	16%	21%	18%
	Feeding tube	2%	0%	1%
MG-ADL total score, mean (SD)	Patient-assessed	7.5 (4.6)	7.7 (3.8)	7.5 (4.3)
MG-ADL total score, mean (SD)	Physician-assessed	8.2 (4.9)	7.9 (3.6)	8.1 (4.5)
Difference	Physician − patient	0.8 (2.3)	0.2 (2.1)	0.6 (2.3)

Abbreviations: AChR, acetylcholine receptor antibody; IVIg, intravenous immunoglobulins; MG-ADL, Myasthenia Gravis Activity of Daily Living score; MGFA classification, Myasthenia Gravis Foundation of America Classification; MuSK, muscle-specific kinase antibody; SD, standard deviation.
[a]The following treatments are not reported due to a response rate of 0%: cyclosporine, tacrolimus, cyclophosphamide, SCIg, and regular use of a breathing mask.

TABLE 2 MG-ADL results by responder.

MG-ADL score	Patient ($N = 137$)	Physician ($N = 137$)	Difference ($N = 137$)
Mean	7.5	8.1	0.6
SD	4.3	4.5	2.3
Q1	4	5	0
Q3	10	11	2

Abbreviations: MG-ADL, Myasthenia Gravis Activity of Daily Living score; Q1, first quartile; Q3, third quartile; SD, standard deviation.

in Italy and Germany. For the MG-ADL individual items, concordance was moderate to substantial for the health change and country/physician subgroups.

4 | DISCUSSION

This study has demonstrated excellent concordance between self and physician assessment of the MG-ADL. The results are consistent across almost all MG-ADL items, indicating that

self-assessment of the MG-ADL is a valid alternative for physician assessment.

The mean MG-ADL score difference between the two assessments was small and physicians estimated that overall symptom severity was similar to or slightly worse than that of the patients themselves. Considering that the minimal important difference (MID) for the MG-ADL has been estimated to be 2 points,[14] the mean difference of 0.6 established in this study can be considered small, indicating excellent overall agreement. Moreover, the mean difference between patient- and physician-assessed MG-ADL scores was consistent across various factors, including countries/physicians, sex, thymectomy status, and antibody status. However, the difference between patient and physician assessment was larger than the MID in 20% of patients (28 of 137), 18 of whom indicated that they had a change in health status in the time between the patient and the physician assessment. In those cases, the difference was reassuringly larger than among those patients who did not experience a change. Larger differences than the MID were also observed in patients who were more severely affected by MG (MGFA class IV), and in patients with four or more comorbidities. Physicians rated symptom burden overall as slightly more severe than their patients, and this finding was stronger with increasing disease severity (MGFA class IVa and IVb). This is consistent with a review by Rand and Caiels of the use of proxy evidence compared with self-report,[15] which showed that physician-rated quality of life was lower than that rated by patients in many studies, which could be the effect of coping mechanisms seen in patients.

TABLE 3 Gwet's AC and ICC for item level and MG-ADL score.

MG-ADL items	Gwet's AC (*P* value)
Talking	0.66 (<0.0001)
Chewing	0.77 (<0.0001)
Swallowing	0.66 (<0.0001)
Breathing	0.73 (<0.0001)
Brush teeth or comb hair	0.58 (<0.0001)
Rise from a chair	0.69 (<0.0001)
Double vision	0.74 (<0.0001)
Eyelid droop	0.46 (<0.0001)
	ICC (95% CI)
MG-ADL total score	0.94 (0.89–0.95)

Abbreviations: CI, confidence interval; Gwet's AC, Gwet's agreement coefficient; ICC, intraclass correlation coefficient; MG-ADL, Myasthenia Gravis Activity of Daily Living score.

The data for this analysis came from two countries, and the German sample was almost twice the size of the Italian sample. The smaller sample size in Italy did not lead to a significantly lower ICC than that in Germany. This is noteworthy, as one previous study found that the concordance of results in samples containing fewer than 50 pairs may be low.[16] A recent "change in health status" was reported twice as often by Italian compared to German patients (21% vs 9%), which may have been the result of slightly longer periods between assessments, or a difference in participants' interpretation of "health change."

Considering the profile of the nine patients removed from the data set, no specific subgroup seemed to have difficulty with self-reporting. Some differences regarding treatments and occurrence of comorbidities were observed between the two countries, but they were not considered significant by the treating physicians and were not expected to have any influence on the results.

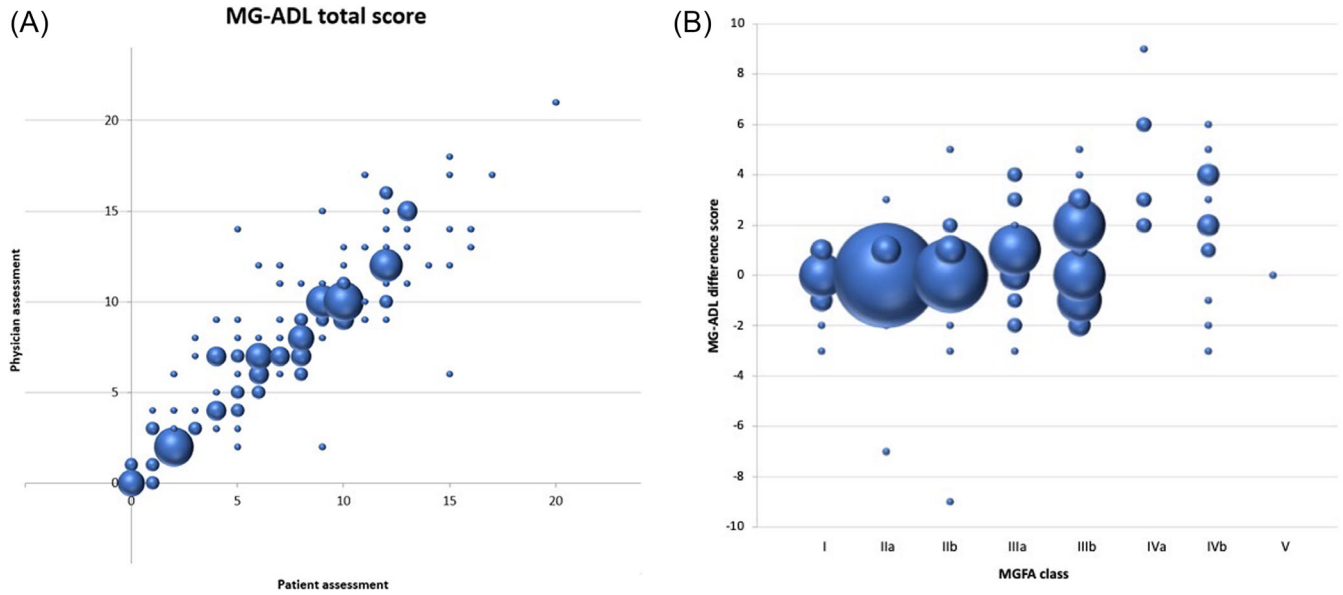

FIGURE 1 (A) Patient vs physician MG-ADL total score and (B) MG-ADL difference score (physician − patient) by current MGFA classification. Bubble size represents the number of observations. The distribution of bubbles across the diagonal axis visualizes the concordance between the proxy and self-assessed scores across all severities. MG-ADL, Myasthenia Gravis Activities of Daily Living scale; MGFA, Myasthenia Gravis Foundation of America clinical classification.

TABLE 4 MG-ADL difference scores (physician − patient) by patient characteristic.[a]

Difference in item score or in total score	Overall	Age (years)			Sex		MGFA class							Thymectomy		AChR		Number of comorbidities			
		18-34	35-54	55+	Male	Female	I	IIa	IIb	IIIa	IIIb	IVa	IVb	Yes	No	Neg.	Pos.	1	2	3	≥4
Talking	0.04	−0.06	0.05	0.05	0.08	0.01	**−0.14**	**0.04**	**0.00**	**0.05**	**−0.07**	**0.29**	**0.50**	0.06	0.04	−0.03	0.06	0.00	−0.13	−0.03	0.17
Chewing	0.10	0.11	0.10	0.10	0.04	0.14	**0.07**	**0.00**	**−0.05**	**0.05**	**0.13**	**0.86**	**0.21**	0.11	0.10	0.00	0.14	**−0.07**	**0.03**	**0.06**	**0.21**
Swallowing	0.15	0.17	0.13	0.15	0.14	0.15	**0.07**	**−0.04**	**−0.05**	**0.14**	**0.27**	**0.86**	**0.29**	0.17	0.11	0.06	0.17	**0.07**	**−0.13**	**0.12**	**0.33**
Breathing	0.04	0.06	−0.08	0.10	0.10	0.00	0.00	0.04	−0.14	−0.05	0.07	0.71	0.14	−0.02	0.10	−0.06	0.07	0.00	0.00	−0.06	0.14
Brush teeth/comb hair	0.22	0.22	0.21	0.23	0.20	0.24	**−0.07**	**0.00**	**0.24**	**0.29**	**0.30**	**0.86**	**0.29**	0.24	0.21	−0.09	0.28	0.14	0.16	0.24	0.26
Rise from a chair	−0.01	0.06	0.10	−0.08	−0.04	0.01	−0.21	0.00	0.00	0.10	−0.07	0.00	0.07	−0.05	0.01	0.00	0.03	−0.07	−0.06	0.06	0.00
Double vision	−0.01	0.11	−0.05	−0.03	0.04	−0.05	−0.07	−0.08	−0.05	0.00	−0.03	0.14	0.07	−0.05	0.01	−0.06	−0.03	0.14	−0.03	−0.15	0.03
Eyelid droop	0.07	**0.39**	**0.15**	**−0.04**	−0.02	0.13	**0.00**	**−0.20**	**−0.05**	**0.10**	**0.13**	**0.71**	**0.43**	0.16	0.01	0.00	0.15	**0.07**	**−0.13**	**−0.06**	**0.26**
Total score	0.60	1.06	0.62	0.49	0.53	0.64	**−0.36**	**−0.24**	**−0.10**	**0.67**	**0.73**	**4.43**	**2.00**	0.63	0.61	**−0.13**	**0.88**	**0.29**	**−0.28**	**0.18**	**1.40**

Abbreviations: AChR, acetylcholine receptor antibody; MG-ADL, Myasthenia Gravis Activity of Daily Living score; MGFA classification, Myasthenia Gravis Foundation of America Classification.

[a]Bold values are statistically significant ($P < 0.05$).

As patients most severely affected by MG (MGFA class IV and V) had a lower prevalence, the highest severity stratum had a lower sample size compared with the other two strata. Possible bias may have occurred in the classification of patients as mild, moderate, or severe, based on their most recent MGFA class rather than a reclassification immediately before participating. Patients may have been incorrectly categorized because of recent worsening of the disease or effective treatment. However, this is unlikely, because the MGFA classifies patients primarily based on symptom localization, which does not evolve rapidly despite the fluctuating nature of MG regarding symptom severity.[14]

We found several differences at the item level between physician and patient assessments. According to Von Essen, a higher degree of correspondence is expected for objective externally observable domains[15,17,18]; therefore, a lower correspondence is expected in domains that cannot be observed in clinical practice. In our study, we found the opposite: there was low concordance on the item *eyelid droop*, but excellent agreement was found in symptoms such as *dysphagia* and *chewing problems*, which are arguably more difficult for a physician to observe in a clinical setting. It is likely that the physicians observed the eyelid droop themselves while completing the MG-ADL questionnaire, whereas the physician probably asked the patient about dysphagia and chewing, which would result in essentially the same answer as when the patient assessed these two items. Overall, these differences did not markedly affect the overall interrater reliability.

As both countries are in Europe, the results may not be generalizable to MG populations on other continents. However, Lee et al conducted a similar study in South Korea, examining the concordance rate among 40 patients with MG and their treating physicians, and demonstrated comparable conclusions. On an item level, good agreement was observed in the South Korean study, with ICCs ranging from 0.645 to 0.985 ($P < 0.001$), although the ICC is not the preferred statistic for assessing concordance at the item level because of the multidimensional nature of the MG-ADL instrument and MG. Remarkably, excellent concordance (ICC = 0.985, $P < 0.001$) was found for the item *eyelid droop* in the South Korean study, whereas concordance for this item in our study, as assessed by Gwet's AC, was only 0.46 ($P < 0.0001$), indicating moderate agreement.[7]

A study investigating the validity and reliability of the MG-ADL instrument compared the change in total MG-ADL scores between two visits with the treating physician's impression of change measured on a 7-point scale (ranging from markedly improved to markedly worsened). A correlation coefficient of 0.703 ($P < 0.0001$) between the change in the actual MG-ADL total score and the impression of change indicated excellent responsiveness of MG-ADL to clinical changes.[19] Our study showed that experiencing a health change during the time between the self and the proxy assessment lowered the concordance between the physician and the patient in the MG-ADL assessment, which is consistent with the finding that MG-ADL is responsive to health changes.

In conclusion, our results suggest that the MG-ADL instrument can be used by patients to self-assess their symptom burden for

TABLE 5 Gwet's AC and ICC for item level and MG-ADL score by health change and country

	Patient having a recent health change		Country	
Items	**No**	**Yes**	**Germany**	**Italy**
Number	**119**	**18**	**95**	**42**
ICC (95% CI)				
MG-ADL total score	0.94 (0.91-0.96)	0.75 (0.32-0.90)	0.93 (0.89-0.95)	0.92 (0.84-0.95)
Gwet's AC (*P* value)				
Talking	0.69 (<0.0001)	0.43 (0.016)	0.68 (<0.0001)	0.60 (<0.0001)
Chewing	0.77 (<0.0001)	0.80 (<0.0001)	0.75 (<0.0001)	0.82 (<0.0001)
Swallowing	0.68 (<0.0001)	0.57 (0.001)	0.68 (<0.0001)	0.63 (<0.0001)
Breathing	0.74 (<0.0001)	0.61 (0.001)	0.71 (<0.0001)	0.74 (<0.0001)
Brush teeth or comb hair	0.58 (<0.0001)	0.58 (0.002)	0.51 (<0.0001)	0.73 (<0.0001)
Rise from a chair	0.72 (<0.0001)	0.43 (0.025)	0.70 (<0.0001)	0.66 (<0.0001)
Double vision	0.76 (<0.0001)	0.64 (<0.0001)	0.81 (<0.0001)	0.59 (<0.0001)
Eyelid droop	0.49 (<0.0001)	0.27 (0.108)	0.50 (<0.0001)	0.39 (<0.0001)

Abbreviations: Gwet's AC, Gwet's agreement coefficient; ICC, intraclass correlation coefficient; MG-ADL, Myasthenia Gravis Activity of Daily Living score.

clinical evaluation, in addition to routine examinations by neurologists. Therefore, the MG-ADL instrument appears to be suitable for self-administered use in future MG-related clinical practice or in clinical research as a primary or secondary outcome measure.

ACKNOWLEDGMENTS

The authors thank all the study participants for sharing their health-related quality of life data.

FUNDING INFORMATION

The study is funded by argenx BV, Ghent, Belgium.

CONFLICT OF INTEREST STATEMENT

S.D., the principal investigator of the study, N.T., and M.F.J. were commissioned by argenx BV and received honoraria to design the study and analyze and report the data. G.P. and S.P. are employees of argenx BV, the sponsor of the study. F.S. received speaking honoraria and honoraria for the participation on advisory boards from Alexion and argenx BV. R.M. has received speaking honoraria from Biomarin, Alexion and UCB; served on advisory boards for Alexion, argenx BV, and UCB; and received support for congress participation from Merck, Teva, and Biogen. A.M. has received speaker honoraria, consulting fees, or financial research support from Alexion, argenx BV, Grifols, Hormosan, Janssen, Octapharma, and UCB. He serves as chairman of the medical advisory board of the German Myasthenia Gravis Society.

DATA AVAILABILITY STATEMENT

Anonymized, aggregated study data is available upon reasonable request through the corresponding author.

ETHICAL PUBLICATION STATEMENT

We confirm that we have read the journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

ORCID

Sarah Dewilde https://orcid.org/0000-0002-7315-3230
M. F. Janssen https://orcid.org/0000-0001-6602-6949
Nafthali H. Tollenaar https://orcid.org/0000-0001-9271-5959
Rita Frangiamore https://orcid.org/0000-0002-0383-3791
Renato Mantegazza https://orcid.org/0000-0002-9810-5737
Andreas Meisel https://orcid.org/0000-0001-7233-5342
Frauke Stascheit https://orcid.org/0000-0001-5306-7880

REFERENCES

1. Conti-Fine BM, Milani M, Kaminski HJ. Myasthenia gravis: past, present, and future. *J Clin Investig*. 2006;116:2843-2854. doi:10.1172/jci29894
2. Wolfe GI, Herbelin L, Nations SP, Foster B, Bryan WW, Barohn RJ. Myasthenia gravis activities of daily living profile. *Neurology*. 1999;52:1487-1489. doi:10.1212/wnl.52.7.1487
3. 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;65:630-639. doi:10.1002/mus.27476
4. Thomsen JLS, Andersen H. Outcome measures in clinical trials of patients with myasthenia gravis. *Front Neurol*. 2020;11:596382. doi:10.3389/fneur.2020.596382
5. Gilhus NE, Tzartos S, Evoli A, Palace J, Burns TM, Verschuuren JJGM. Myasthenia gravis. *Nat Rev Dis Primers*. 2019;5:30. doi:10.1038/s41572-019-0079-y
6. Melzer N, Ruck T, Fuhr P, et al. Clinical features, pathogenesis, and treatment of myasthenia gravis: a supplement to the guidelines of the German Neurological Society. *J Neurol*. 2016;263:1473-1494. doi:10.1007/s00415-016-8045-z
7. Lee HL, Min JH, Seok JM, et al. Physician- and self-assessed Myasthenia Gravis Activities of Daily Living score. *Muscle Nerve*. 2018;57:419-422. doi:10.1002/mus.25764
8. Temel GO, Erdoğan S. Determining the sample size in agreement studies. *Marmara Med J*. 2017;30:101-112.
9. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol*. 2008;61:29-48. doi:10.1348/000711006x126600
10. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol*. 1990;43:551-558. doi:10.1016/0895-4356(90)90159-m

11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174. doi:10.2307/2529310
12. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155-163. doi:10.1016/j.jcm.2016.02.012
13. Stata statistical software. Release 16. College Station, Texas: StataCorp 2019.
14. Barnett C, Herbelin L, Dimachkie MM, Barohn RJ. Measuring clinical treatment response in myasthenia gravis. *Neurol Clin*. 2018;36:339-353. doi:10.1016/j.ncl.2018.01.006
15. Rand S, Caiels J. *Using proxies to assess quality of life: A review of the issues and challenges*. Discussion paper 2899. Quality and Outcomes of person-centred care policy Research Unit (QORU). Editor. Canterbury: University of Kent; 2015. Retrieved from: https://www.pssru.ac.uk/pub/4980.pdf
16. Sneeuw KCA, Sprangers MAG, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease. *J Clin Epidemiol*. 2002;55:1130-1143. doi:10.1016/S0895-4356(02)00479-1
17. von Essen L. Proxy ratings of patient quality of life-factors related to patient-proxy agreement. *Acta Oncol*. 2004;43:229-234. doi:10.1080/02841860410029357
18. Diez Porras L, Homedes C, Alberti MA, Velez Santamaria V, Casasnovas C. Quality of life in myasthenia gravis and correlation of MG-QOL15 with other functional scales. *J Clin Med*. 2022;11:2189.
19. Muppidi S, Wolfe GI, Conaway M, Burns TM. MG-ADL: still a relevant outcome measure. *Muscle Nerve*. 2011;44:727-731. doi:10.1002/mus.22140

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Dewilde S, Janssen MF, Tollenaar NH, et al. Concordance between patient- and physician-reported Myasthenia Gravis Activities of Daily Living (MG-ADL) scores. *Muscle & Nerve*. 2023;68(1):65-72. doi:10.1002/mus.27837