

# Validation of the MG-DIS: a disability assessment for myasthenia gravis

Alberto Raggi<sup>1</sup> · Matilde Leonardi<sup>1</sup> · Silvia Schiavolin<sup>1</sup> · Carlo Antozzi<sup>2</sup> · Greta Brenna<sup>2</sup> · Lorenzo Maggi<sup>2</sup> · Renato Mantegazza<sup>2</sup>

Received: 9 November 2015 / Revised: 12 February 2016 / Accepted: 15 February 2016 / Published online: 1 March 2016  
© Springer-Verlag Berlin Heidelberg 2016

**Abstract** This paper is aimed to present the validation of the myasthenia gravis disability assessment (MG-DIS), a MG-specific patient-reported disability outcome measure. Consecutive MG patients were enrolled, followed-up for 12 months and administered the SF-36, the WHO disability assessment schedule (WHODAS 2.0) and the preliminary 31-item MG-DIS addressing impairments and activity limitations. Factor structure and metric properties were assessed. In total, 109 patients were enrolled: 76 were females, mean age 50, mean MG duration 10.4 years, 86 were AChR-positive. The MG-DIS was reduced to 20 items, explaining 70.6 % of the original questionnaire variance, four subscales (generalized impairment-related problems; bulbar function-related problems; mental health and fatigue-related problems; vision-related problems) and an overall disability index. The MG-DIS has good metric properties (Cronbach's alpha ranging between .808 and .930), is stable, showed to be more sensitive than the WHODAS 2.0 and SF-36 to detect group differences and longitudinal changes and was well correlated with the MG-composite (.642). The MG-DIS includes items representing ocular, generalized, bulbar and respiratory symptoms, and

is therefore well-built around MG-specific features. MG-DIS can be used in clinical trials as well as in observational or epidemiological studies to characterize patients' disability level and address the amount of improvement in disability. Further studies are needed to explore the possibility of a shorter disability scale.

**Keywords** Myasthenia gravis · Disability evaluation · Factor analysis · Validation study · Patient-reported outcome

## Introduction

Myasthenia gravis (MG) is an autoimmune disease caused by autoantibodies to the post-synaptic site of the neuromuscular junction, whose main features are fluctuating muscle weakness and fatigability [1]. Typically, MG affects women with a bimodal period of onset, around 20–40 and 60–80 years of age [2, 3]. Staging of MG recognizes different forms, depending on the severity of weakness and fatigability and on muscle involvement (ocular, generalized, bulbar or respiratory). As shown in a literature review, outcome measures for MG are mostly focused on clinical manifestations that are deemed to be immediately evident and relevant for clinicians dealing with MG patients [4]. However, due to the unique characteristics of MG (i.e. the fact that symptoms are evident to the patient and can fluctuate during the day), it is important that patient-reported outcomes measures (PROMs) are incorporated in patient's assessment. This was also one of the recommendations for clinical trials in MG, which highlighted the need to incorporate subject responses into trial outcome measures [5]. PROMs are increasingly used in clinical trials and in descriptive studies, as they complete

**Electronic supplementary material** The online version of this article (doi:10.1007/s00415-016-8072-9) contains supplementary material, which is available to authorized users.

✉ Alberto Raggi  
araggi@istituto-besta.it

<sup>1</sup> Neurology, Public Health and Disability Unit, Neurological Institute C. Besta IRCCS Foundation, Via Celoria 11, 20133 Milan, Italy

<sup>2</sup> Department of Neuroimmunology and Neuromuscular Disorders, Neurological Institute C. Besta IRCCS Foundation, Milan, Italy

the assessment on the benefit of intervention, which is essential to provide evidence of the impact over patients in terms of health status, health-related quality of life (HRQoL) or disability [6–8]. HRQoL was well correlated with different measures of disability in patients with MG, including patient-reported [9, 10] and physician-reported measures [11, 12]: with regard to PROM-referred correlations, it has been shown that the two measures do not address the same construct and therefore should not be considered as transposable [9].

MG clinical fluctuations, together with the need of taking drugs for a prolonged period of time, make patients' perspective very important to investigate MG clinical course and measure treatments' effectiveness. Disease-specific measures addressing HRQoL, such as the MG-specific quality of life questionnaire (MG-QOL) and its shorter 15-item version (MG-QOL-15) have been developed [13, 14]. These include a mixture of items relating to symptoms (e.g. trouble using eyes) and limitations of daily activities (e.g. trouble driving). However, no commonly used patient-reported outcome measure that addresses disability and that focuses on symptoms (including both those directly and those not directly associated to MG) and on the difficulties in daily activities complained by patients exists. The result of this is that studies aiming to address disability in MG patients have to rely on non-specific measures, such as the modified Rankin Scale [11], the Incapacity Status Scale [15] or the WHO disability assessment schedule (WHODAS 2.0) [9, 10]. For this reason, we developed, tested and validated the myasthenia gravis disability assessment (MG-DIS): aim of this paper is to present the validation study of this questionnaire.

## Methods

The process of development of the MG-DIS has been described in a previous publication [16]. In brief, based on a review of the literature, content of outcome measures used in MG were linked to the categories of the International Classification of Functioning, Disability and Health (ICF) [17]: a total of 13 ICF categories were identified in this way. These categories were compared to those previously identified as relevant to describe disability in MG patients [18], to form a longer list of categories (55 in total): these categories were used to interview a group of patients and exclude ICF categories that were not systematically reported by patients. In this way, 42 ICF categories were retained and used to develop the 44-items preliminary version of the MG-DIS questionnaire used for this validation study: of these 44 items, 31 addressed impairments and limitations in performing daily activities, while 13 addressed contextual factors that might act as facilitators or

barriers in the perspective of the patient. In this study, we focused on the 31 items addressing impairments and limitations only.

## Patients and setting

Patients' enrolment was carried out between April 2013 and April 2014 at the Neurological Institute C. Besta, and follow-up was concluded on April 2015. We included patients aged 18 or older, with a diagnosis of MG based on clinical data and one of the following: positive response to acetylcholinesterase inhibitors, positive acetylcholine receptor (AChR) or muscle-specific tyrosine kinase (MuSK) antibody assay, decrement of more than 10 % in the amplitude of the compound muscle action potential on repetitive nerve stimulation, or increased jitter on single-fiber electromyography. In double negative (DN) patients, the diagnosis of MG was confirmed by neurophysiologic investigations. Patients with comorbidity to other diseases with autoimmune features and those with severe respiratory impairment requiring mechanical ventilation were excluded.

Participation to the study was on a voluntary basis, and patients were enrolled on the occasion of hospital admission or outpatient visit. The project was approved by the Institute's Ethical Committee, and all patients signed an informed consent form prior to their inclusion in the study.

## Measures

The case report forms included a demographic section, a section on patients' clinical features and a section with PROMs.

Demographic information included age, gender, marital status, years of education and higher level of education, employment status. Patient's clinical status was classified with our Besta Neurological Institute rating scale for myasthenia gravis (INCB-MG) [19, 20], that provides a numerical score for muscle strength and fatigability, and describes MG muscle involvement in four areas (ocular, generalized, bulbar and respiratory). The scale was initially reported in 1988 [19], used routinely at our Institution for clinical assessment of MG patients, and its formal validation has been recently released [20]. Muscle fatigability was measured as time (s) of upper and lower limbs maintained outstretched: the amount of time in seconds that patients were able to keep the position (up to 120 s) was recorded. In addition to INCB-MG, the MG-composite was also administered to patients to address clinical severity [21]. It is a ten-item assessment addressing MG symptoms referred to ocular, generalized and bulbar muscle groups: scores range between 0 and 50, with higher scores indicating worse clinical status. Age at disease onset,

autoantibody profile (AChR, MuSK or DN), previous thymectomy and ongoing medical therapy were collected from clinical records: for each category (steroids, acetylcholinesterase inhibitors and immunosuppressants) daily dose was recorded. It was also recorded whether patients were treated with plasma exchange or intravenous immunoglobulins (IVIG) in the previous month.

PROMs included the MG-DIS and two additional questionnaires: the WHODAS 2.0 [22], and the Medical Outcomes Study 36-item short-form health survey (SF-36) [23].

The 31 items addressing impairments and limitations were defined from the perspective of the patients, who had to rate how much of a problem (or how much of a difficulty) they had in the last thirty days: response option was on a 1–5 scale varying between “no problem/no difficulty” and “complete problem/complete difficulty”. The time-frame of the previous 30 days is expected to allow patients to account for MG fluctuations but, at the same time, it is close enough to avoid reporting bias.

The WHODAS 2.0 is a 36-items disability assessment tool that examines difficulties due to a health condition: patients have to answer questions regarding how much difficulty, due to their health condition, they had in the last 30 days. Six subscales (understanding and communicating; getting around; self-care; getting along with people; life activities, divided into household and work; participation in society) and a total score are available, with range on a 0–100 scale, with higher scores reflecting greater disability.

The SF-36 measures eight health concepts dealing with both physical and mental status. Two main scores are available: physical and mental composite scores (PCS and MCS). SF-36 scores range on a 0–100 scale, with higher scores reflecting better HRQoL, while PCS and MCS are norm-based scores (mean 50 and standard deviation 10, with higher scores reflecting better HRQoL). Items are referred to the past 4 weeks.

## Data analysis

Continuous variables were reported as means and standard deviations (SD) or means and 95 % confidence intervals (95 % CI), discrete variables as medians and interquartile range (IQR). Data were analysed with SPSS 19.0.

### Factor structure and item reduction

Prior to carrying out the exploratory factor analysis (EFA), symmetry indexes were evaluated: items that were clearly asymmetric (symmetry index  $\geq 2.58$ ) were eliminated from the dataset. We also looked at the intercorrelation between items and removed those with an overall inter-item correlation index below 0.300. We also removed those items

that showed correlation indexes  $>0.800$  with at least three other items (i.e. more than 10 % of the total number of items): in this way, multicollinearity or singularity problems are expected to be avoided [24]. Suitability of data for factor analysis was assessed with Bartlett’s test of sphericity (BTS; adequate if  $P < .05$ ) [24], which tests whether the variables are sufficiently correlated one to the other, and with Kaiser–Meyer–Olkin measure of sampling adequacy (KMO) [25]: KMO values between 0.5 and 0.7 are mediocre, values between 0.7 and 0.8 are good, values between 0.8 and 0.9 are great and values above 0.9 are superb [24].

Direct oblimin rotation was used to extract data, as we reasonably expect that factors might be correlated. Kaiser criterion for extraction was used [26]. We chose this criterion since we began with 31 variables, but some were expected to be dropped out before extraction on the basis of intercorrelations: a confirmation on its adequacy is tested on the basis of the analysis of communalities that should, on average, be higher than 0.700.

Three steps to item reduction were carried out. First, items that did not load into any factor (i.e. with factor loadings  $<.40$ ) were deleted as they give no contribution to questionnaire’s structure. Second, items loading into multiple factors (i.e. with factor loadings  $\geq .40$  into more than one factor) were deleted as they would determine high instability in MG-DIS factor structure. Third, we looked at inter-item correlation matrix among items loading into the same factors and addressed those with correlation  $\geq .800$ . These items are deemed to have similar information content, and were therefore candidate for being deleted: the one with the higher loading was retained.

Model fit was tested with ratio between Chi Square and  $df$  (good if  $\leq 3$ ) and root mean square error of approximation (RMSEA; good if  $\leq 0.08$ ) [25]. Score of MG-DIS scales were calculated on a 0–100 basis, with higher scores reflecting higher disability. To get to the 0–100 scale, raw scores were developed on the basis of response to the items, and a linear transformation was used: transformed score = (raw score – min score)/(max score – min score)  $\times$  100. For example, a raw score of 8 on a hypothetical four-item scale (i.e. as if all items were score as “a little problem”), would be  $(8 - 4) / (20 - 4) \times 100$ , i.e.  $4/16 \times 100 = 25$ .

### Internal consistency

Internal consistency was assessed using Cronbach’s alpha coefficient, item-total correlation after correcting for overlap (i.e. removing the item from the total score) and the average inter-item correlation. Scales were considered to have good reliability if Cronbach’s alpha  $\geq .70$  [27], if item-total correlation indexes were  $\geq 0.40$  and average inter-item correlations were comprised between 0.30 and

0.70 [28]. Internal consistency analyses were repeated also for the subgroup of patients with positive AChR antibody assay and for non-AChR (i.e. MuSK and DN).

#### *Construct validity*

Construct validity was tested in two ways. First, by correlating MG-DIS scores to muscle fatigability indexes for upper and lower limbs and to MG-composite. It is expected that: correlations between MG-DIS scores and muscle fatigability indexes are significant and inverse (i.e. the higher the value of MG-DIS scales, the lower the time in second with limb outstretched); that correlations between MG-DIS scores and MG-composite are significant and direct (i.e. to higher MG-DIS scores correspond higher MG-composite scores); that muscle fatigability indexes and MG-composite are better correlated to the MG-DIS than to WHODAS 2.0 summary score, SF-36 PCS and MCS. Second, by assessing differences in MG-DIS scales between patients in remission or with ocular symptoms and those with generalized or bulbar symptoms: the latter stages are in fact those in which it is expected to find a strong impact of the disease over a wider set of daily activities. Differences were assessed using Mann–Whitney *U* test, and magnitude of difference using Hedges' *g* effect size (ES) measure:  $ES \geq 0.80$  are considered to reflect large differences. It is expected that the MG-DIS performs better than the WHODAS 2.0 and the SF-36, in terms magnitude of group differences. The ability of the MG-DIS compared to that of the WHODAS 2.0 and SF-36 PCS and MCS in discriminating patients with bulbar or generalized symptoms was also assessed calculating the area under the receiver operating characteristic curve (ROC): it is expected that the value of the area of MG-DIS is wider than those of WHODAS 2.0 and SF-36 PCS and MCS.

#### *Stability*

The stability of items and factors was assessed with a short-term test–retest analysis: as items have to be rated considering the previous 30 days, the 5-days retest makes the two periods overlap at 80 %. Stability was assessed using item-by-item and factor-by-factor Spearman's correlations: it is expected that correlation coefficients are  $\geq .400$ .

#### *Sensitivity to change*

Patients completed their regular follow-up at the institute: on occasion of the first outpatient examination or admission to ward, those that were enrolled completed a second administration of questionnaires. Baseline differences between study completers and drop-out were assessed using Mann–Whitney *U* test.

Clinical change was measured by means of the INCB-MG scale [19, 20]. Patients were classified as: (a) improved, if follow-up score was at least 60 % better (in case patients remained in the same muscle involvement category, i.e. ocular, generalized or bulbar), or if patients moved from a worse towards a better category; (b) worsened, if follow-up score was at least 60 % worse (in case patients remained in the same muscle involvement category), or if patients moved from a better towards a worse category; (c) unchanged, if baseline and follow-up clinical manifestation were similar and patients remained in the same muscle involvement category. We did not consider, for the purpose of this analysis, any reduction or increase in medication, as suggested in the guideline provided by the post-intervention status evaluation [29]. There are two reasons for this: first, clinical change was defined with a different assessment procedure and, second, our interest was to address the effect of MG on patients' disability. Therefore, a substantial reduction in medication, but stable MG profile, would not be consistent with the aim of addressing the reduction of disability perceived by patients. Repeated measure analysis were carried out using Wilcoxon signed-rank test and Cohen's *d* ES as a measure of magnitude of change: it is expected that patients that improved report a lower (i.e. better) score at MG-DIS and that ES is  $\geq 0.80$ ; it is expected that patients that worsened report a higher (i.e. worse) score at MG-DIS and that ES is  $\geq 0.80$ ; it is expected that unchanged patients report similar scores and that ES is  $< 0.30$ . It is also expected that the MG-DIS performed better than the WHODAS and the SF-36, i.e. that ES referred to the group of patients that improved and worsened are higher than those observed for WHODAS and SF-36. The mean and 95 % CI between baseline and follow-up score difference was also computed to provide an indication of how much of a change in MG-DIS is expectable in improved and worsened patients.

In addition to this, the difference between baseline and follow-up scores at MG-DIS, WHODAS 2.0 and MG-composite [21] were calculated: correlations between these delta scores were calculated, with the hypothesis that the delta of MG-composite was better correlated to the delta of MG-DIS than to the delta of WHODAS 2.0 scores.

## **Results**

A total of 114 MG patients were eligible: of them, two refused to participate, three were unable to participate for health reasons unrelated to MG. Thus, 109 patients, 76 females, aged 22–80 (mean 50, SD 15.6) were enrolled: Table 1 provides main clinical and demographic features. Average WHODAS 2.0 total score was 25.4 (SD 18.9),

**Table 1** Sample characteristics

Age	50.0 (15.6)	
Gender		
Females	76 (69.7 %)	
Males	33 (30.3 %)	
Marital status		
Never married	31 (28.4 %)	
Married/cohabitating	71 (65.1 %)	
Separated/divorced/widowed	7 (6.5 %)	
Years of education	11.9 (3.8)	
Education level		
Up to secondary	46 (42.2 %)	
High	47 (43.1 %)	
Academic	16 (14.7 %)	
Employment status		
Employed/student	44 (40.4 %)	
Homemaker	6 (5.5 %)	
Retired	36 (33 %)	
Unemployed	12 (11 %)	
Unemployed due to health reasons <sup>a</sup>	11 (10.2 %)	
Setting		
Outpatients	77 (70.6 %)	
Inpatients	27 (24.8 %)	
Day-hospital treatment	5 (4.6 %)	
Age at the onset (years) <sup>b</sup>	39.6 (17.1)	
Disease duration (years)	10.4 (10)	
MG type		
Remission	33 (30.3 %)	
Ocular MG	22 (20.2 %)	
Generalized MG	34 (31.2 %)	
Bulbar MG	20 (18.3 %)	
Autoantibody profile		
AChR	86 (78.9 %)	
MuSK	13 (11.9 %)	
Double negative	10 (9.2 %)	
Muscle fatigability, s		
Upper limbs	100.4 (33.6)	
Lower limbs	84.9 (41.5)	
MG-composite	3.4 (4.6)	
Previous thymectomy	56 (51.3 %)	
Distance from thymectomy (years)	11.8 (10.7)	
Therapy <sup>c</sup>		Average daily dose (mg)
Steroids	71 (65.5 %)	21.5 (15.3)
Acetylcholinesterase inhibitor	70 (64.2 %)	295 (118.4)
Immunosuppressants		
Azathioprine	38 (34.9 %)	136.8 (43.4)
Mycophenolate mofetil	5 (4.6 %)	1200 (237.9)
Methotrexate	4 (3.7 %)	2.8 (1.5)
Cyclophosphamide	3 (2.8 %)	83.3 (57.7)

Categorical variables are reported as frequencies and percentages, continuous variables as means and SD

<sup>a</sup> One patient was on maternal leave

<sup>b</sup> Age at onset was prior to 40 years in 59 patients (54.1 %) and after 40 years in 50 patients (45.9 %)

<sup>c</sup> 18 patients received intravenous immunoglobulins or underwent plasmapheresis within the previous month



average PCS was 39.3 (SD 10.8) and average MCS was 44.0 (SD 12.0).

### Factor analysis and internal consistency

Based on correlation analysis, we excluded two items (memory problems; problems with allergies or infections) as their overall inter-item correlation index was  $<0.300$ , and three items as they showed correlation indexes  $>0.800$  with three or more items and were asymmetric (problems with muscle weakness: its highest correlation was .889 with “muscle fatigability”). Problems with household chores: its highest correlation was .815 with “shopping for daily needs”. Problems with work activities: its highest correlation was .817 with “muscle fatigability”). Therefore, factor analysis was carried out over 26 items. Please refer to supplementary material for information on excluded items.

KMO was .890 and Bartlett’s test was significant at  $P < .001$  level (Chi Square: 2072.6;  $df$  325), therefore the sample is adequate for size, and variables are sufficiently correlated one to the other. The average communality was .713, therefore Kaiser criterion was adequate.

A four-factor solution explained 70.6 % of the total variance of original questionnaire with an adequate fit: Chi Square/ $df$  ratio was 1.61, RMSEA was 0.075. No items had factor loadings below 0.400; four items loaded into two factors, and were excluded; two items had correlation  $>.800$  with another item from the same factor, and were excluded due to low factor loading. Therefore, the final MG-DIS was composed of 20 items. Please refer to supplementary material for detailed factor analysis data.

Table 2 reports the factor structure and internal consistency, as well as item-level properties of the MG-DIS. The factors had average inter-item correlations between 0.465 and 0.641, average item-total correlations between 0.596 and 0.726, and Cronbach’s alpha coefficients between 0.808 and 0.911. Considering the entire section, alpha was 0.930, average inter-item correlation was .400 and average item-total correlation was .609. Deleting item 4.3 (difficulties driving) would make alpha of factor 4 to increase from .846 to .848, but such a change would not significantly impact on the factor’s internal consistency and we preferred to maintain the item due to its relevance: being unable to drive, in fact, determines important problems with personal autonomy and mobility. Similar results were found when scores of patients positive to AChR and those non positive to AChR were tested for reliability analyses: inter-item correlations varied between 0.419 and 0.689, average item-total correlations between 0.462 and 0.756, and Cronbach’s alpha coefficients between 0.711 and 0.922. Considering the entire section, alpha was 0.945 in AChR patients and 0.914 in non-AChR patients; average

inter-item correlation was .452 non-AChR patients and 0.366 in non-AChR patients; average item-total correlation was .658 non-AChR patients and 0.586 in non-AChR patients. Please refer to supplementary material for detailed reliability analysis data.

Mean and SD for MG-DIS subscales and overall disability index are reported in Table 3.

### Construct validity

Table 3 also reports correlations between MG-DIS scales, upper and lower limb muscle fatigability indexes and MG-composite: as expected, the MG-DIS was inversely correlated with muscle fatigability indexes and directly correlated with MG-composite. MG-DIS correlations with muscle resistance indexes were slightly superior ( $-0.581$ ;  $-0.558$ ) than those of WHODAS 2.0 ( $-0.467$ ;  $-0.546$ ) and SF-36 PCS (0.440; 0.536), and clearly superior than those of SF-36 MCS (0.210; 0.232). Similarly, the MG-composite had higher correlation with the MG-DIS (0.642) than with the WHODAS 2.0 (0.492) and SF-36 PCS and MCS ( $-0.422$ ;  $-0.238$ ).

Table 4 reports group differences. The MG-DIS discriminated well patients in remission or with ocular symptoms from those with generalized or bulbar symptoms: the ES was indicative of a large difference and was superior than that observed for the WHODAS 2.0. The same was for the subscales: with the exception of the “mental health and fatigue-related problems” scale, the other scales discriminated very well between the two groups of patients. This result is further on reinforced by the analysis of the area under the ROC, that was higher for MG-DIS (.809) compared to WHODAS 2.0 (.739) and SF-36 scores (.706; .627).

### Stability

Test–retest analysis was carried out on 21 patients (18 females, mean age 43.2). Spearman’s correlations were all significant ( $P < .001$ ) and all higher than .800 when factors were taken into account (see Table 3). At the level of single items, correlation coefficients ranged between .574 and .960 (refer to supplementary material for full test–retest correlation matrix).

### Sensitivity to change

Complete follow-up information was available for 75 patients, with average follow-up duration of 241 days. There were no baseline differences between study completers and drop-outs for age, years of education, MG duration, age at onset, upper and lower muscle fatigability indexes, WHODAS 2.0, SF-36 PCS and MCS and MG-DIS

**Table 2** Factor structure and internal consistency of MG-DIS, and item-level properties

In the last 30 days...	Factor loadings	Item mean ± SD	Item-total correlation	Alpha if item excluded
Factor 1. Generalized impairment-related problems [alpha = .911; inter-item <i>R</i> = .593; average item-total <i>R</i> = .726, Eigenvalue: 11.95 (46 % of total variance)]				
(1.1) ...did you have difficulties in dressing, for example putting on or taking off socks and footwear, skirts or trousers, shirts, hats, gloves?	.914	1.5 ± 0.9	.724	.879
(1.2) ...did you have difficulties in washing and drying yourself, for example, having a shower, washing your face or hairs?	.825	1.6 ± 1.1	.768	.873
(1.3) ... did you have difficulties in picking up small objects (e.g. coins from a table), in using cutlery or scissors?	.746	1.4 ± 0.8	.679	.885
(1.4) ... did you have difficulties in cutting food and moving it to the mouth?	.741	1.3 ± 0.6	.739	.887
(1.5) ... did you have difficulties in shopping for daily needs?	.713	1.8 ± 1.1	.774	.873
(1.6) ... did you have difficulties in opening bottles and cans to drink a beverage?	.659	1.6 ± 0.9	.749	.876
(1.7) ... did you have problems due to muscle fatigability?	.430	2.4 ± 1.3	.648	.896
Factor 2. Bulbar function-related problems [alpha = .868; inter-item <i>R</i> = .563; average item-total <i>R</i> = .694, Eigenvalue: 2.62 (10.1 % of total variance)]				
(2.1) ... did you have difficulties in eating (i.e. chewing and ingesting food)?	.901	1.7 ± 1.0	.821	.805
(2.2) ... did you have difficulties in chewing or swallowing (liquids or food)? Did you have problems with excessive salivation or vomiting?	.885	1.9 ± 1.1	.781	.815
(2.3) ... did you have difficulties in drinking (i.e. swallowing a beverage)?	.863	1.4 ± 0.8	.691	.841
(2.4) ... did you have difficulties in talking with other persons or have difficulties in expressing yourself?	.681	1.6 ± 0.9	.589	.862
(2.5) ... did you have voice problems, such as hyper nasality, hoarseness or being voiceless?	.644	1.7 ± 1.0	.589	.864
Factor 3. Mental health and fatigue-related problems [alpha = .808; inter-item <i>R</i> = .465; average item-total <i>R</i> = .596, Eigenvalue: 1.79 (8.4 % of total variance)]				
(3.1) ... did you feel anxious sad, low or angry?	.795	2.3 ± 1.0	.606	.768
(3.2) ... did you have bodily ache, such as headache, back, neck or stomach pain?	.628	2.6 ± 1.1	.545	.787
(3.3) ... did you have problems with sleeping, such as falling asleep, waking up during the night, or not feeling refreshed in the morning?	.625	2.5 ± 1.2	.664	.749
(3.4) ... did you feel without energy or listless, had lack of appetite or had difficulties in controlling your reactions?	.482	2.3 ± 1.1	.667	.748
(3.5) ... did you have difficulties in breathing or felt breathless?	.403	1.8 ± 1.0	.497	.799
Factor 4. Vision-related problems (alpha = .846; inter-item <i>R</i> = .641; average item-total <i>R</i> = .698, Eigenvalue: 1.59 (6.1 % of total variance)]				
(4.1) ... did you have difficulties in reading?	.903	1.9 ± 1.1	.776	.688
(4.2) ... did you have sight problems, such as blurred or double vision, or had difficulties with seeing near or distant objects?	.844	2.2 ± 1.2	.706	.737
(4.3) ... did you have difficulties in driving a car or a motorbike?	.505	2.0 ± 1.5	.612	.848

total scores. Of the 34 that did not complete the follow up, there were no information for 11 of them (ten were not regularly followed-up in the institute, one died for reasons unrelated to MG) while clinical follow-up was available for 23. Of them, 12 were on remission, and therefore did not come to visit by 12 months: clinical follow-up showed that all of them were still on remission. The remaining 11 were not interested in completing the follow-up: of them, six were stable (two with ocular and four with generalized symptoms), four improved (one from bulbar to ocular symptoms, two from bulbar to generalized symptoms, one from generalized symptom to pharmacological remission) and one worsened (from ocular to generalized symptoms).

Table 5 shows longitudinal differences across unchanged, worsened and improved patients. No differences were found for unchanged patients, in none of the outcome measures: the majority of them (42.5 %) were on remission or had ocular symptoms (22.5 %). With regard to worsened patients, significant differences were found for all MG-DIS scales, with ES higher than 0.8 (except for “mental health and fatigue-related problems” scale), and thus indicative of a large difference. The SF-36 PCS scale also showed a significant and moderate longitudinal difference. On the contrary, no significant difference was found for WHO-DAS 2.0 despite ES was higher than 0.80. Mean difference between baseline and follow-up scores for worsened

**Table 3** Mean and SD for MG-DIS, stability analysis (short-term test–retest), and correlations with upper and lower limbs resistance indexes

Scale	Mean (SD)	Test–retest <sup>a</sup>	Correlations		
			Upper limbs	Lower limbs	MG-composite
MG-DIS overall disability	22.0 (17.0)	.943	–.581	–.558	.642
Generalized impairment-related problems	16.4 (19.5)	.961	–.552	–.553	.537
Bulbar function-related problems.	16.5 (19.4)	.919	–.434	–.442	.696
Mental health and fatigue-related problems	32.4 (20.7)	.858	–.448	–.434	.429
Vision-related problems	26.4 (27.2)	.939	–.339	–.414	.403

<sup>a</sup> At the level of single items, test–retest correlation ranged between .574 and .960. All correlations were significant at  $P < .001$

**Table 4** Discriminant validity

	Mean (95 % CI)		<i>P</i> value	Effect size	AUROC
	Remission/ocular	Generalized/bulbar			
WHODAS 2.0	18.3 (13.6–23.0)	32.7 (27.9–37.5)	.001	0.76	.739
SF-36 PCS	43.0 (40.3–45.8)	35.5 (32.7–38.3)	.001	0.69	.706
SF-36 MCS	46.4 (43.1–49.7)	41.5 (38.3–44.6)	.023	0.41	.627
MG-DIS overall disability	13.2 (10.1–16.4)	30.8 (26.2–35.5)	.001	1.04	.809
Generalized impairment-related problems	8.3 (5.3–11.2)	24.7 (18.5–30.9)	.001	0.84	.760
Bulbar function-related problems.	5.9 (3.6–8.2)	27.2 (21.3–33.1)	.001	1.10	.801
Mental health and fatigue-related problems	25.4 (20.2–30.7)	39.5 (34.1–44.9)	.001	0.73	.703
Vision-related problems	16.5 (10.3–22.7)	36.4 (28.9–44.0)	.001	0.87	.713

Significance was set at  $P < .006$  after Bonferroni correction. Effect size  $>0.80$  indicate large differences

WHODAS 2.0, WHO disability assessment schedule, second version; PCS, Physical Composite Score; MCS, Mental Composite Score; MG-DIS, MG-specific disability assessment; AUROC, area under the receiver operating curve

patients was 16.6 (95 % CI 5.0–28.2). The majority of these patients had ocular (27.3 %) or generalized symptoms (36.4 %). With regard to improved patients, only MG-DIS scales showed significant differences, with ES that for MG-DIS global score and for “mental health and fatigue-related problems” scale were indicative of a large longitudinal difference. Mean difference between baseline and follow-up scores for improved patients was 17.7 (95 % CI 9.2–26.2). The majority of these patients had generalized (54.2 %) or bulbar symptoms (33.3 %).

Finally, change in MG-composite was much more strongly correlated to change in MG-DIS (.688,  $P < .001$ ) than to change in WHODAS 2.0 (.379,  $P = .001$ ).

## Discussion

With this paper we tested and validated the MG-DIS, a MG-specific PROM for the assessment of disability. It is composed of an overall disability index and of four sub-scores: generalized impairment-related problems; bulbar function-related problems; mental health and fatigue-related problems; vision-

related problems. The MG-DIS index showed to have good metric properties: it had good to excellent Cronbach’s alpha coefficients, also when tested separately for AChR and non-AChR patients, and showed to be stable in the short-term retest analysis. Compared to the reference questionnaires, WHODAS 2.0 and SF-36, it showed a stronger discriminative power in distinguishing patients with more severe forms of MG, stronger correlations with muscle fatigability indexes and MG-composite, and showed higher sensitivity in capturing improvement over time.

The most innovative aspect of MG-DIS is the inclusion of items that enable to address a wide variety of difficulties and symptoms representative of the different degree of muscle involvement and MG severity, thus representing ocular (e.g. difficulties in reading), generalized (e.g. difficulties in dressing, in washing and drying yourself) and bulbar problems (e.g. difficulties in eating/drinking, voice problems) as well as difficulties with respiration or sensation of being breathless. A note on respiratory symptom is needed. Patients requiring intubation, which represent a minority of MG patients, were excluded from our sample. As reported in a recent a



**Table 5** Sensitivity to change analysis

	Mean (95 % CI)		P	ES
	Baseline	Follow-up		
<b>Worsened (N = 11)</b>				
FU duration: 271 days				
MG duration: 12.9 years				
WHODAS	20.6 (12.4–28.7)	33.6 (20.9–46.3)	.010	1.06
SF-36 PCS	38.8 (33.0–44.7)	32.3 (24.9–39.6)	.008	0.75
SF-36 MCS	45.8 (37.5–54.2)	43.0 (36.7–49.3)	.477	0.22
MG-DIS	18.8 (9.9–27.7)	35.4 (26.7–44.0)	.005	1.25
Generalized impairment-related problems	13.6 (5.6–21.9)	27.9 (15.5–40.3)	.006	1.16
Bulbar function-related problems.	18.2 (2.1–34.2)	41.8 (25.9–57.7)	.032	0.99
Mental health and fatigue-related problems	28.2 (14.9–41.5)	40.5 (32.3–48.6)	.005	0.62
Vision-related problems	15.9 (3.3–28.5)	33.3 (23.0–43.6)	.003	0.93
<b>Unchanged (N = 40)</b>				
FU duration: 250 days				
MG duration: 11.0 years				
WHODAS	20.0 (14.7–25.3)	19.1 (13.8–24.4)	.191	0.05
SF-36 PCS	42.9 (39.4–46.4)	43.3 (40.2–46.4)	.610	0.04
SF-36 MCS	45.8 (42.1–49.5)	46.9 (43.6–50.2)	.368	0.08
MG-DIS	16.1 (11.8–20.4)	16.9 (12.2–21.6)	.994	0.06
Generalized impairment-related problems	12.2 (6.8–17.5)	13.5 (7.9–19.1)	.518	0.08
Bulbar function-related problems.	8.9 (4.8–13.0)	9.6 (4.3–15.0)	.954	0.05
Mental health and fatigue-related problems	25.5 (19.7–31.3)	26.4 (20.3–32.4)	.260	0.05
Vision-related problems	21.5 (14.4–28.5)	21.2 (15.0–27.5)	.987	0.01
<b>Improved (N = 24)</b>				
FU duration: 212 days				
MG duration: 5.2 years				
WHODAS	33.2 (26.4–40.1)	27.1 (20.5–33.7)	.014	0.38
SF-36 PCS	36.4 (32.6–40.3)	37.5 (33.3–41.7)	.391	0.12
SF-36 MCS	41.3 (36.7–46.1)	46.1 (42.1–50.1)	.081	0.43
MG-DIS	29.6 (21.5–37.6)	11.8 (8.8–14.9)	<.001	0.94
Generalized impairment-related problems	23.7 (14.0–33.4)	9.4 (6.2–12.6)	.003	0.62
Bulbar function-related problems.	24.2 (13.8–24.5)	7.3 (4.3–10.3)	.006	0.69
Mental health and fatigue-related problems	39.6 (30.8–48.4)	20.2 (14.8–25.6)	<.001	0.93
Vision-related problems	35.4 (22.5–48.4)	11.1 (8.1–14.2)	.001	0.79

For each subset of patients, significance was set at  $P < .006$  after Bonferroni correction. Effect size  $>0.80$  indicate large differences

WHODAS 2.0, WHO disability assessment schedule, second version; PCS, Physical Composite Score; MCS, Mental Composite Score; MG-DIS, MG-specific disability assessment

retrospective study based on a cohort of 677 MG patients treated in our institute [30], need for intubation was detected as “maximal worsening” in 6.2 % of patients, and in no cases at “last observation”; in a Spanish cohort of 648 patients need for intubation was reported in 4.9 % over 13 years [31]. Reduced vital capacity was instead observed in 39 % of patients in a large cohort evaluated between 1966 and 2000 [32]. However, respiratory symptoms in MG patients do not necessarily mean need for intubation or reduced vital capacity and, although very

relevant, have been poorly investigated. In fact, respiratory symptoms are addressed in physician-reported measures, such as the MG composite [21], while are generally excluded from patient-reported outcomes (e.g. the MG-QOL [13] and the MG questionnaire [33]): up to now only the MG activities of daily living profile [34] included one item on respiration. The fact that MG-DIS also includes one item on respiration makes it possible to detect the higher disability severity in this subgroup of patients.

The MG-DIS also includes an item referred to bodily pain, which is not considered a MG-specific symptom, although it has been shown to affect patients' daily life and HRQoL [35]. Pain has been generally considered of interest in MG after thymectomy [36, 37] but, when specifically asked to patients, pain was reported between 50.6 % [38] and 60.8 % [18] of cases: therefore the inclusion of pain in a measure of disability is thought to provide a strong contribution to the understanding of the outcome. Other non-MG specific issues are on the contrary included in the MG-QOL-15 [14], such as feelings of depression and difficulties in driving, while sleeping problems are addressed only in the full 60-items version of the MG-QOL [13].

Compared to the other two questionnaires employed in this study, the MG-DIS summary score showed higher discriminative power: ES were indicative of a large difference between patients with generalised/bulbar symptoms vs. ocular/remission, which was instead moderate for WHODAS 2.0 and SF-36, and this result was further on confirmed by the area under the ROC. Moreover, the MG-DIS showed a stronger association with upper and lower limb muscle fatigability and with an established clinical severity scale such as the MG-composite, with correlation coefficients ( $r = .642$ ,  $P < .001$ ) that suggest association, but not overlap. In our opinion, the reason for this is that the MG-DIS covers all areas of possible muscle involvement, which enabled to address problems that might be relevant for patients with generalised and, more specifically, bulbar symptoms. Our sample comprised a high number of patients with bulbar symptomatology (18.3 %), which was higher compared to our recently published retrospective where the corresponding figure was 7.8 % [30].

Longitudinal analyses also showed the superiority of the MD-DIS compared to the WHODAS 2.0 and SF-36, which is also reinforced by the correlations between the delta of MG-composite and the delta of MG-DIS and WHODAS 2.0. These analyses show that the variation described by clinicians with the MG-composite is consistent, but not overlapping ( $r = .688$ ,  $P < .001$ ), with that reported by patients with the MG-DIS, while the level of correspondence with the WHODAS 2.0 is clearly lower. The majority of unchanged patients were on pharmacological remission, or had ocular symptoms only, at baseline. These patients are not likely to change over time from a clinical point of view, which on one side might have diminished the capacity of the MG-DIS to detect changes, but strongly confirmed its stability. With regard to clinically improved patients, the ES of MG-DIS was 0.94, indicative of large difference, while those of WHODAS 2.0 and SF-36 were comprised between 0.12 and 0.43, and therefore indicative of mild or no differences. With regard to worsened patients, again the MG-DIS showed to be superior: the ES

of MG-DIS was 1.25 and was statistically significant ( $P = .005$  at Wilcoxon signed-rank test); the ES was lower for SF-36 (between 0.22 and 0.75), and again high for WHODAS 2.0 (ES = 1.06), but statistically not significant. It has however to be noted that only eleven patients (14.7 % of those with complete follow-up) worsened and the paucity of numbers might have played a role. The fact that there is a discrepancy between the ES values and the non-significance of Wilcoxon signed-rank test is likely to confirm this hypothesis. Moreover, as the sample is composed of patients with quite a long experience of being affected by MG (average disease duration was 10.4 years), it is possible that disease duration also had an effect on their evaluation: those that worsened had a longer history of MG (12.9 years on average) than those that improved (5.2 years). It is therefore possible that patients that worsened were more "acquainted" to the fluctuations of MG, with a diminished effect on the way they perceive their general HRQoL and disability: this further on confirms the importance of measuring disability with an assessment instrument specifically designed over MG features, like the MG-DIS. The relevance of fluctuations for PROMs measures in MG was already reported by Barnett and colleagues as strictly connected to fatigability and to the difficulties in daily activities [39]. Such an hypothesis, however, should be tested in future studies focused on the effect of history of fluctuations over patients perception of their health state. In other words, would patients with a history of frequent fluctuation report their disability level differently from those with a more stable disease?

One separate comment is deserved on the elimination of the item referred to muscle weakness. Although clinical evaluation distinguishes muscle weakness from fatigability, our data show that patients reported the two items as substantially overlapping, as the bivariate correlation was .889. This does not mean that the content of the two items are perceived as equal, but that patients tended to report a very similar evaluation, i.e. if muscle fatigability was perceived as a severe problem (score 4), it was very unlikely that muscle weakness was perceived with no problem (score 1); rather, it was very likely that it was perceived as a moderate to complete problem (score between 3 and 5). Therefore, maintaining weakness item would have produced a redundant over-reporting of a problem that was already described: this marks the difference between patients' and physicians' reports.

Another separate comment is deserved on the features of MG-DIS items. Some of them are quite narrow: for example, all those referred to eating and drinking. These two activities are underlined by generalized (cutting food; opening bottles) and bulbar involvement (chewing and ingesting food; swallowing a beverage) that were perceived as distinct by patients at the first step of MG-DIS

development: such a distinction was also confirmed—at least for the Italian language—in the validation process. On the contrary, other are quite broad: the most outstanding example of this is the item referred to “dressing” (i.e. putting on or taking off socks and footwear, skirts or trousers, shirts, hats, gloves), that merges different contents and also implies the use of different muscle groups. The content of the item is the activity of dressing which is a clearly identified activity despite variation due to gender (i.e. wearing skirts), season (i.e. the use of coats in cold periods) and contexts (i.e. the use of formal or informal clothes). It has moreover to be noted that the version herein presented was validated in Italian language, and the item included in the manuscript did not undergo a formal process of translation and back-translation (the Italian version is included in supplementary materials).

Some limitations have to be taken into account. First, clinical follow-up was available for 23 of the 34 patients that dropped out: of them, 18 were stable (12 were on pharmacological remission, six were symptomatic), four improved and one worsened. In consideration of the good correspondence between the clinicians’ and patients’ evaluations, it is possible to presume that no or minor differences might have occurred for stable patients, and some kind of improvement is likely to be expected for improved patients. The group of patients that were not regularly followed-up was mostly composed of patients with generalised or bulbar symptoms (9 out of 11): these patients might have improved over time or be stable, i.e. the case of refractory MG [40] and the impact of this on the analysis of sensitivity to change is impossible to predict. For these reasons, the drop-out rate at follow-up suggests a cautious interpretation of results referred to the longitudinal evaluation. Second, as the sample is composed of patients with quite a long experience of being affected by MG, it is possible that disease duration also had an effect on their evaluation. In fact, those that worsened had a longer history of MG than those that improved, and it is therefore reasonable to think that they were more “acquainted” to the fluctuations of MG. Third, the sample was mostly composed of AChR-positive patients, whose disease course might be quite different compared to MuSK-positive and DN ones and this might had an impact on the validity of our questionnaire: however, reliability analyses were tested separately for AChR-positive and non-AChR-positive patients, showing that MG-DIS was reliable in both groups. Fourth, the high prevalence of patients on remission and the relevant prevalence of them among stable patients might have affected the capacity of the MG-DIS to detect changes.

In conclusion, we presented the results of the validation of the MG-DIS, a 20-items patient-reported MG-specific disability assessment instrument covering all MG-related

symptoms that demonstrated good metric properties. Further studies are needed to explore the possibility of a shorter disability scale, and to address issues arose in the present study with a larger multicenter study. The MG-DIS can be used in clinical trials as well as in observational or epidemiological studies to characterize patients’ disability level and to address the amount of improvement in disability jointly with clinical improvement. Joining clinicians’ and patients’ outcomes reporting will be of particular relevance for the characterization of patients with refractory MG, as it will enable to comprehensively address benefits of new treatments.

**Acknowledgments** This study was supported by Besta Institute Foundation.

#### Compliance with ethical standards

**Conflicts of interest** The authors declare that they have no conflict of interest.

**Ethical standards** The study was approved by the Institute’s ethics committee, and was therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

## References

- Conti-Fine BM, Milani M, Kaminski HJ (2006) Myasthenia gravis: past, present, and future. *J Clin Invest* 116:2843–2854
- Mantegazza R, Baggi F, Antozzi C, Confalonieri P, Morandi L, Bernasconi P, Andreetta F, Simoncini O, Campanella A, Beghi E, Cornelio F (2003) Myasthenia gravis (MG): epidemiological data and prognostic factors. *Ann NY Acad Sci* 998:413–423
- Vincent A, Palace J, Hilton-Jones D (2001) Myasthenia gravis. *Lancet* 357:2122–2128
- Burns TM (2010) History of outcome measures for myasthenia gravis. *Muscle Nerve* 42:5–13
- Benatar M, Sanders DB, Burns TM, Cutter GR, Guptill JT, Baggi F, Kaminski HJ, Mantegazza R, Meriggioli MN, Quan J, Wolfe GI, Task Force on MG Study Design of the Medical Scientific Advisory Board of the Myasthenia Gravis Foundation of America (2012) Recommendations for myasthenia gravis clinical trials. *Muscle Nerve* 45:909–917
- Fitzpatrick R, Davey C, Buxton M, Jones DR (2001) Criteria for assessing patient based outcome measures for use in clinical trials. In: Stevens A, Abrams K, Brazier J, Fitzpatrick R, Lilford R (eds) *The advanced handbook of methods in evidence based healthcare*. Sage, London, pp 181–194
- Sanders C, Egger M, Donovan J, Tallon D, Frankel S (1998) Reporting on quality of life in randomised controlled trials: bibliographic study. *BMJ* 317:1191–1194
- Fitzpatrick R, Davey C, Buxton MJ, Jones DR (1998) Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 2:1–74
- Leonardi M, Raggi A, Antozzi C, Confalonieri P, Maggi L, Cornelio F, Mantegazza R (2010) The relationship between health, disability and quality of life in myasthenia gravis: results from an Italian study. *J Neurol* 257:98–102

10. Raggi A, Leonardi M, Antozzi C, Confalonieri P, Maggi L, Cornelio F, Mantegazza R (2010) Concordance between severity of disease, disability and health-related quality of life in myasthenia gravis. *Neurol Sci* 31:41–45
11. Farrugia ME, Carmichael C, Cupka BJ, Warder J, Brennan KM, Burns TM (2014) The modified rankin scale to assess disability in myasthenia gravis: comparing with other tools. *Muscle Nerve* 50:501–507
12. Barnett C, Katzberg H, Nabavi M, Bril V (2012) The quantitative myasthenia gravis score: comparison with clinical, electrophysiological, and laboratory markers. *J Clin Neuromuscul Dis* 13:201–205
13. Mullins LL, Carpentier MY, Paul RH, Sanders DB, Muscle Study Group (2008) Disease-specific measure of quality of life for myasthenia gravis. *Muscle Nerve* 38:947–956
14. Burns TM, Conaway MR, Cutter GR, Sanders DB, Muscle Study Group (2008) Less is more, or almost as much: a 15-item quality-of-life instrument for myasthenia gravis. *Muscle Nerve* 38:957–963
15. Chen YT, Chang Y, Chiu HC, Yeh JH (2011) Psychosocial aspects in myasthenic patients treated by plasmapheresis. *J Neurol* 258:1240–1246
16. Raggi A, Schiavolin S, Leonardi M, Antozzi C, Baggi F, Maggi L, Mantegazza R (2014) Development of the MG-DIS: an ICF-based disability assessment instrument for myasthenia gravis. *Disabil Rehabil* 36:546–555
17. World Health Organization (2001) International classification of functioning, disability and health (ICF). WHO, Geneva
18. Leonardi M, Raggi A, Antozzi C, Confalonieri P, Maggi L, Cornelio F, Mantegazza R (2009) Identification of international classification of functioning, disability and health relevant categories to describe functioning and disability of patients with myasthenia gravis. *Disabil Rehabil* 31:2041–2046
19. Mantegazza R, Antozzi C, Peluchetti D, Sghirlanzoni A, Cornelio F (1988) Azathioprine as a single drug or in combination with steroids in the treatment of myasthenia gravis. *J Neurol* 235:449–453
20. Antozzi C, Brenna G, Baggi F, Camera G, Maggi L, Rezzani C, Montomoli C, Mantegazza R (2016) Validation of the besta neurological institute rating scale for myasthenia gravis. *Muscle Nerve* 53:32–37. doi:10.1002/mus.24911
21. Burns TM, Conaway M, Sanders DB, MG Composite and MG-QOL15 Study Group (2010) The MG composite: a valid and reliable tool for myasthenia gravis. *Neurology* 74:1434–1440
22. Garin O, Ayuso-Mateos JL, Almansa J, Nieto M, Chatterji S, Vilagut G, Alonso J, Cieza A, Svestkova O, Burger H, Racca V, Francescutti C, Vieta E, Kostanjsek N, Raggi A, Leonardi M, Ferrer M, MHADIE consortium (2010) Validation of the “World Health Organization disability assessment schedule, WHODAS-2” in patients with chronic diseases. *Health Qual Life Outcomes* 8:51
23. Ware JE (2000) SF-36 health survey update. *Spine* 25:3130–3139
24. Field A (2009) *Discovering statistics using SPSS*, 3rd edn. Sage, London
25. Hu L, Bentler PM (1994) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model* 6:1–55
26. Kaiser HF (1974) An index of factor simplicity. *Psychometrika* 39:31–36
27. Bland JM, Altman DG (1997) Cronbach’s alpha. *BMJ* 314:572
28. Nunnally JC, Bernstein I (1994) *Psychometric theory*, 3rd edn. McGraw-Hill, New York
29. Jaretzki A 3rd, Barohn RJ, Ernstoff RM, Kaminski HJ, Keesey JC, Penn AS, Sanders DB (2000) Myasthenia gravis: recommendations for clinical research standards. Task Force of the Medical Scientific Advisory Board of the Myasthenia Gravis Foundation of America. *Neurology* 55:16–23
30. Baggi F, Andretta F, Maggi L, Confalonieri P, Morandi L, Salerno F, Bernasconi P, Montomoli C, Barberis M, Mantegazza R, Antozzi C (2013) Complete stable remission and autoantibody specificity in myasthenia gravis. *Neurology* 80:188–195
31. Ramos-Fransi A, Rojas-García R, Segovia S, Márquez-Infante C, Pardo J, Coll-Cantí J, Jericó I, Illa I, Myasthenia NMD-ES Study Group (2015) Myasthenia gravis: descriptive analysis of life-threatening events in a recent nationwide registry. *Eur J Neurol* 22:1056–1061
32. Grob D, Brunner N, Namba T, Pagala M (2008) Lifetime course of myasthenia gravis. *Muscle Nerve* 37:141–149
33. Padua L, Evoli A, Aprile I, Caliandro P, Batocchi AP, Punzi C, Mazza S, Padua R, Tonali P (2002) Myasthenia gravis outcome measure: development and validation of a disease-specific self-administered questionnaire. *Neurol Sci* 23:59–68
34. Muppidi S, Wolfe GI, Conaway M, Burns TM, MG Composite and MG-QOL15 Study Group (2011) MG-ADL: still a relevant outcome measure. *Muscle Nerve* 44:727–731
35. Scott KR, Kothari MJ (2006) Self-reported pain affects quality of life in myasthenia gravis. *J Clin Neuromuscul Dis* 7:110–114
36. Su C, Su Y, Chou CW, Liu W, Zou J, Luo H, Chen Z (2012) Intravenous flurbiprofen for post-thymectomy pain relief in patients with myasthenia gravis. *J Cardiothorac Surg* 7:98
37. Nilsson E, Perntunen K, Kalso E (1997) Intrathecal morphine for post-sternotomy pain in patients with myasthenia gravis: effects on respiratory function. *Acta Anaesthesiol Scand* 41:549–556
38. Guy-Coichard C, Nguyen DT, Delorme T, Boureau F (2008) Pain in hereditary neuromuscular disorders and myasthenia gravis: a national survey of frequency, characteristics, and impact. *J Pain Symptom Manag* 35:40–50
39. Barnett C, Bril V, Kapral M, Kulkarni A, Davis AM (2014) A conceptual framework for evaluating impairments in myasthenia gravis. *PLoS One* 9:e98089
40. Silvestri NJ, Wolfe GI (2014) Treatment-refractory myasthenia gravis. *J Clin Neuromuscul Dis* 15:167–178