

## Concise report

## Assessment of disability in idiopathic inflammatory myopathy: a call for linearity

Minoesch Min<sup>1,\*</sup>, Anne W. Walter <sup>1,\*</sup>, Johan Lim <sup>1</sup>, Filip Eftimov<sup>1</sup>, Camiel Verhamme<sup>1</sup>, Marianne de Visser<sup>1</sup>, Ivo N. van Schaik<sup>1,3</sup>, Rohit Aggarwal<sup>4</sup>, Rob J. de Haan<sup>2</sup>, Anneke J. van der Kooi<sup>1,\*</sup> and Joost Raaphorst<sup>1,\*</sup>, the Dutch Myositis Network

## Abstract

**Objectives.** To evaluate the clinimetric properties of the Academic Medical Centre Disability Score (ALDS) in patients with idiopathic inflammatory myopathy (IIM).

**Methods.** We used prospectively collected data of IIM patients who completed a phase-2 study with first-line IVIG monotherapy. The ALDS is a patient-reported questionnaire which contains 25 items relevant for disability in myositis. ALDS and all core set measures (CSMs) for myositis [including HAQ-Disability Index (HAQ-DI)] were evaluated at baseline and 9 weeks follow-up. In addition, the 2016 ACR/EULAR myositis response criteria outcome called Total Improvement Score (TIS) was evaluated at 9 weeks. We examined floor/ceiling effects, reliability and construct validity of the ALDS. To examine known-group validity, ALDS change scores over time were compared with TIS and physician impression of clinical response.

**Results.** Nineteen patients with IIM [median age 59 years, 12 (63%) female] were enrolled. At baseline, ALDS showed a median score of 65.4 (IQR 58.2–73.5), good Cronbach's alpha ( $\alpha=0.84$ ) and a small ceiling effect (11%). Construct validity was confirmed by moderate to strong correlations between ALDS and HAQ-DI [ $r_s=-0.57$  (baseline);  $-0.86$  (follow-up)]. ALDS change score correlated with TIS ( $r_s=0.70$ ), discriminated between responders and non-responders ( $TIS \geq 40$ ;  $P=0.001$ ), between groups based on physician impression of clinical response ( $P=0.03$ ), and detected deterioration.

**Conclusion.** The ALDS showed promising clinimetric properties and detected relevant changes in disability in patients with myositis. These results warrant further investigations.

**Key words:** myositis, idiopathic inflammatory myopathies, disability score, clinimetric evaluation, ALDS

## Rheumatology Key messages

- The ALDS has promising clinimetric properties in newly diagnosed patients with idiopathic inflammatory myopathy.
- The ALDS detects changes in disability, including deterioration.

## Introduction

Assessment of treatment effects in patients with idiopathic inflammatory myopathy (IIM) is challenging [1, 2].

Recently, the ACR and EULAR approved validated response criteria called Total Improvement Score (TIS) to assess treatment effect in IIM patients. TIS is a composite measure combining disability, different measures of disease and patient's and physician's impressions [3, 4]. Despite the increasing use of TIS in clinical practice and research, it is important to develop validated outcome

<sup>1</sup>Department of Neurology, Amsterdam UMC, University of Amsterdam, Amsterdam Neuroscience, <sup>2</sup>Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam UMC, University of Amsterdam, Amsterdam, <sup>3</sup>Spaarne Gasthuis, Haarlem, The Netherlands and <sup>4</sup>Department of Rheumatology and Clinical Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

Submitted 15 September 2021; accepted 29 November 2021

Correspondence to: Joost Raaphorst, PO Box 22660, 1100DD Amsterdam, The Netherlands. E-mail: j.raaphorst@amsterdamumc.nl

\*Minoesch Min, Anne W. Walter, Anneke J. van der Kooi and Joost Raaphorst contributed equally to this study.

measures that specifically represent (single) domains. Functioning, captured by outcome measures that assess disability, is one of the most important outcome measures from a patient's perspective [1, 5].

Disability in the TIS composite score is measured by the HAQ Disability Index (HAQ-DI) which has been adopted from rheumatoid arthritis and has not been fully validated in myositis. The HAQ-DI is an ordinal scale that has some methodological shortcomings in terms of clinimetrics [6].

Linear scales (as opposed to ordinal scales, such as the Myositis Activity Profile questionnaire) are based on Item Response Theory with one of the unique advantages that similar change scores on different points on a scale represent similar changes in function. The Academic Medical Center Linear Disability Score (ALDS) is a linear scale and enables the use of a subset of items from an item bank to obtain a detailed picture of disability. ALDS has been validated in patients with stroke, Parkinson's disease and rheumatoid arthritis [7–10]. We aimed to evaluate the clinimetric properties of the ALDS in IIM patients.

## Methods

### Patients

We included adult patients with newly diagnosed, treatment-naïve, biopsy-proven IIM with a disease duration of <9 months and a minimal disability of at least 10% loss on manual muscle testing (MMT)-score. We used data of all patients included in a phase-2 open-label study investigating IVIG as first-line monotherapy (IMMEDIATE study) [11].

### Patient data

Data were obtained at baseline (upon inclusion, before treatment) and after 9 weeks of treatment and included demographic and disease variables [11], ALDS questionnaire, core set measures (CSMs) and EQ5D5L (a generic quality of life measure) [12]. The study was conducted in accordance with the Declaration of Helsinki and the ethical committee of the Amsterdam Medical Centre approved the study protocol. All patients gave written informed consent.

The CSMs include the Physician Global Activity, Patient Global Activity, MMT, HAQ-DI and Muscle Enzyme and Extramuscular Activity. The TIS is obtained by adding up all the weighted individual improvement scores of the six core measurements and ranges from zero to 100 [3]. The EQ5D5L consists of five dimensions (e.g. mobility) with five response levels, leading to a five-digit index that can be transformed to a total quality of life score [13]. The physician impression of clinical response (PICR), assigned by the treating physician, was based on all available information except TIS, and included the categories slightly, moderately and markedly improved and slightly, moderately and markedly deteriorated.

### ALDS questionnaire

The ALDS is a generic item bank of 73 ADL items ordered from simple to complex activities. A panel of three physicians with myositis expertise and a patient (myositis representative of the Dutch Patient Association of Neuromuscular Diseases) selected 25 relevant items with a wide range of complexity (Supplementary Data S1, available at *Rheumatology* online). The properties of the scale, in terms of the Item Response Theory (IRT) and the linearity of the scale, ensure that 20–25 items are sufficient to provide robust ALDS scores [10]. Patients rated their ability to carry out activities. Answer options were: (i) 'yes', (ii) 'yes with difficulty', (iii) 'no' and (iv) 'do not know'. To calculate ALDS scores, i and ii were both scored as 'yes' [7, 14]. Original units of the ALDS scale are (logistic) regression coefficients, expressed in logits [7]. To facilitate interpretation of results, logit scores were linearly transformed into ALDS values between 0 and 90, with lower scores indicating more disability.

### Clinimetric evaluation of the ALDS

We used baseline and follow-up data to examine clinimetric properties of the ALDS:

- i. Floor and ceiling effects were calculated by the number (%) of patients with minimal and maximal ALDS scores.
- ii. Reliability was expressed in terms of homogeneity, referring to the statistical coherence of scale items (baseline data).
- iii. Construct validity (referring to whether ALDS measures the intended construct) was examined by correlating ALDS (change) scores with the different CSMs, TIS and EQ5D5L; correlation coefficients were considered weak ( $r \leq 0.30$ ), moderate ( $r = 0.40$ – $0.60$ ) or strong ( $r \geq 0.70$ ) [15].
- iv. Known-group validity was calculated using a change score (follow-up *minus* baseline ALDS score). We compared ALDS change scores:
  - v. Between responders and non-responders based on  $TIS \geq 40$  (at least moderate improvement) and between patients with and without minimal improvement ( $TIS \geq 20$ ) [3]. For these comparisons effect sizes (Hedge's  $g$ ) were calculated using logit scores.
  - vi. Between groups based on TIS scores: no (<20), minimal (20–39), moderate (40–59) and major improvement ( $\geq 60$ ) [3].
  - vii. Between different levels of physician impression of clinical response (PICR): slightly, moderately and markedly improved; and slightly, moderately and markedly deteriorated. For analyses, we used four categories (categories 'slightly' and 'moderately' were combined).

### Statistical analysis

Patient characteristics and outcome measurements were summarized using descriptive statistics.

ALDS scores were calculated using previously published algorithms [7, 14]. Missing items or 'do not know'

responses were discarded [14]. We used original ALDS logits for analyses and linearly transformed ALDS scores for presentation.

Reliability of the ALDS was expressed as Cronbach's  $\alpha$  ( $\geq 0.80$  indicates good homogeneity) [16]. Associations between ALDS (change) scores and CSMs (change) as well as TIS scores were expressed as Spearman correlation coefficients ( $r_s$ ). Between-group differences of ALDS baseline and change scores were analysed using Mann–Whitney  $U$  test or Kruskal–Wallis test, where appropriate. When the Kruskal–Wallis test showed statistically significant score differences across groups, we performed post-hoc pairwise comparisons. A  $P$ -value  $< 0.05$  was considered statistically significant. In view of the explorative nature of this study, we did not correct for multiple comparisons [17].

## Results

### Patient characteristics

Nineteen patients with IIM were included [63% female; median age 59 years (IQR 37–69); median duration between first symptoms and diagnosis was 5 months (IQR 3–6)]. Eight (42%) patients had DM, six (32%) had immune mediated necrotizing myopathy, four (21%) had non-specific or overlap myositis (NM/OM) and one (5%) had antisynthetase syndrome. Ten (53%) patients had myositis-specific antibodies [MSA; MDA-5 ( $n=1$ ), TIF-1 $\gamma$  ( $n=1$ ), NXP-2 ( $n=3$ ), SRP ( $n=1$ ), Jo-1 ( $n=1$ ) and HMGCR ( $n=3$ )], two (11%) patients had myositis-associated antibodies [MAA; Ku ( $n=1$ ), U1RNP ( $n=1$ )] and six (32%) patients were seronegative. Three (16%) patients had a concomitant connective tissue disorder (mixed connective tissue disease, Sjögren's syndrome and systemic sclerosis) and one patient with DM was diagnosed with ovarian cancer three weeks after inclusion.

### Clinimetric evaluation of the ALDS

At baseline, ALDS showed no floor effect; two patients had a maximum score (ceiling = 11%). At follow-up, no patient had a minimum score and seven patients achieved the maximum score (37%). Reliability of ALDS at baseline was considered good (Cronbach's  $\alpha = 0.84$ ). The correlation between ALDS and HAQ-DI was moderate to strong [ $r_s = -0.57$  (baseline) and  $-0.86$  (follow up);  $P < 0.05$ ; Table 1]. In addition, patient and physician global impressions showed moderate correlation with ALDS and correlation was lower with less related constructs such as CK and extra-muscular activity. Muscle strength (MMT) had low correlation at baseline, at follow-up there was a statistically significant moderate correlation with ALDS. The association between ALDS change scores and change in several relevant CSMs such as HAQ-DI, MMT, patient and physician global impressions, as well as the EQ5D5L, was moderately

strong. Moreover, ALDS change correlated strongly with TIS ( $r_s = 0.70$ ,  $P < 0.01$ ). For comparison of clinimetric properties between the ALDS and the HAQ-DI, we show correlations between HAQ-DI and the CSMs, TIS and EQ-5D, respectively in [Supplementary Table S1](#), available at *Rheumatology* online.

### Known-group validity: comparison with responders according to TIS

ALDS change scores discriminated between responders (based on a TIS  $\geq 40$ ;  $n=8$ ) and non-responders ( $n=11$ ): within-group change scores were 18.7 and 0.0 in responders and non-responders, respectively ( $P < 0.01$ ; Hedge's  $g$  effect size  $-1.82$ ; Fig. 1A). Baseline ALDS scores did not differ ( $P = 0.93$ ). ALDS change scores also differed between patients with ( $n=10$ ) and without ( $n=9$ ) minimal improvement (TIS  $\geq 20$ ; within-group change scores 18.5 vs 0.0;  $P < 0.05$ ; Hedge's  $g$  effect size  $-1.12$ ).

### Known-group validity: ALDS in relation to improvement on the TIS

ALDS change scores differed between groups based on TIS categories: no, minimal, moderate and major improvement. Within-group change scores were 0.0,  $-3.2$ , 19.6 and 17.7, respectively ( $P < 0.01$ ; Fig. 1B); baseline ALDS scores did not differ ( $P = 0.19$ ). Post-hoc analysis showed differences between 'no improvement' vs 'major improvement' ( $P = 0.03$ ) and 'no improvement' vs 'moderate improvement' ( $P < 0.01$ ).

### Known-group validity: ALDS in relation to physician impression of clinical response (PICR)

Within-group ALDS change scores differed between groups based on PICR: slightly/moderately deteriorated ( $-0.9$ ), markedly deteriorated ( $-48.3$ ), slightly/moderately improved (9.2) and markedly improved (18.7;  $P = 0.03$ ; Fig. 1C). Baseline ALDS scores did not differ ( $P = 0.28$ ). Post-hoc analysis showed a difference between 'markedly improved' and 'slightly/moderately worse' ( $P = 0.01$ ).

## Discussion

We examined clinimetric properties of a generic linear disability scale in IIM patients treated with monotherapy IVIG. Findings were compared with Total Improvement Score (TIS), individual core set measures (CSMs) and EQ5D5L. ALDS showed promising clinimetric properties: good reliability, no floor effect and a small ceiling effect; moderate to high correlations with HAQ-DI (a measure of disability) and lower correlations with less related constructs. ALDS showed known-group validity and discriminated between different levels of clinical improvement based on TIS and a physician's impression. The moderate to high correlations with a measure of quality of life confirm the importance of disability for patients with myositis.

**TABLE 1** Construct validity of the ALDS: Test scores and Spearman's correlation coefficients of ALDS, CSMs, TIS and EQ5D5L

Outcome measurement	Test score (baseline)	Spearman's correlation with ALDS (baseline)	Test score (follow-up <sup>a</sup> )	Spearman's correlation with ALDS (follow-up)	Spearman's correlation of change scores with ALDS change score
ALDS	65.4 (58.2–73.4)	—	75.6 (65.1–89.1) <sup>b</sup>	—	—
PhGA	3.8 (3.2–4.0)	–0.46*	2.3 (1.0–4.0) <sup>b</sup>	–0.71**	–0.60**
PGA	6.1 (5.3–7.6)	–0.49*	4.6 (2.0–6.6) <sup>b</sup>	–0.43 <sup>†</sup>	–0.64**
MMT	211 (185–225)	0.28	227 (191–241)	0.69**	0.61**
HAQ-DI	2.0 (1.5–2.5)	–0.57*	1.6 (0.8–2.1) <sup>b</sup>	–0.86**	–0.77**
CK	1199 (179–6500)	–0.20	196 (83–3877) <sup>b</sup>	–0.45	–0.29
EMA	2.2 (0.6–3.0)	–0.11	1.0 (0.3–2.3) <sup>b</sup>	0.10	–0.17
TIS			35 (15–53)	0.54 <sup>†</sup>	0.70**
EQ5D5L	0.45 (0.41–0.57)	0.52*	0.60 (0.43–0.78)	0.78**	0.68**

Score values are presented as median (IQR); Spearman's correlation coefficients of ALDS logits with TIS core set measures. \*Significant at the 0.05 level. \*\*Significant at the 0.01 level. <sup>a</sup>Follow-up = nine weeks (or premature ending of participation in the study). <sup>b</sup>Significantly different compared to baseline ( $p < 0.05$ ; Wilcoxon signed rank test). ALDS: Academic medical centre Linear Disability Score; CK: Creatine kinase; EMA: Extramuscular Activity; EQ5D5L (index): a generic quality of life measure; HAQ-DI: HAQ-Disability Index; MMT: Manual Muscle Testing; PGA: Patient Global Assessment; PhGA: Physician Global Assessment; TIS: Total Improvement Score.

The ceiling effect at baseline was larger as expected, possibly due to little disability in some patients [9, 10]. We chose ALDS items with divergent difficulty, but excluded most difficult items (e.g. vacuum a flight of stairs), which could have contributed to the ceiling effect. Another explanation is related to answer options: patients answered many difficult items as 'yes with difficulty', which were analysed as 'yes' (according to the scoring algorithm of ALDS).

Despite some ceiling effect, ALDS provides unique and additional information as compared with HAQ-DI and TIS. Firstly, the linear design of ALDS improves clinical interpretation of scores, as opposed to ordinal scales (HAQ-DI) of which identical score differences may not represent identical clinical changes at another point on the scale [18].

Secondly, ALDS score is a unidimensional and interpretable outcome, opposed to TIS which is a composite of (changes of) CSMs, making it easier to interpret as a supplementary outcome in myositis clinical trials. ALDS also provides a meaningful baseline value, which facilitates interpretation of clinical trial results (variation in baseline disability between groups).

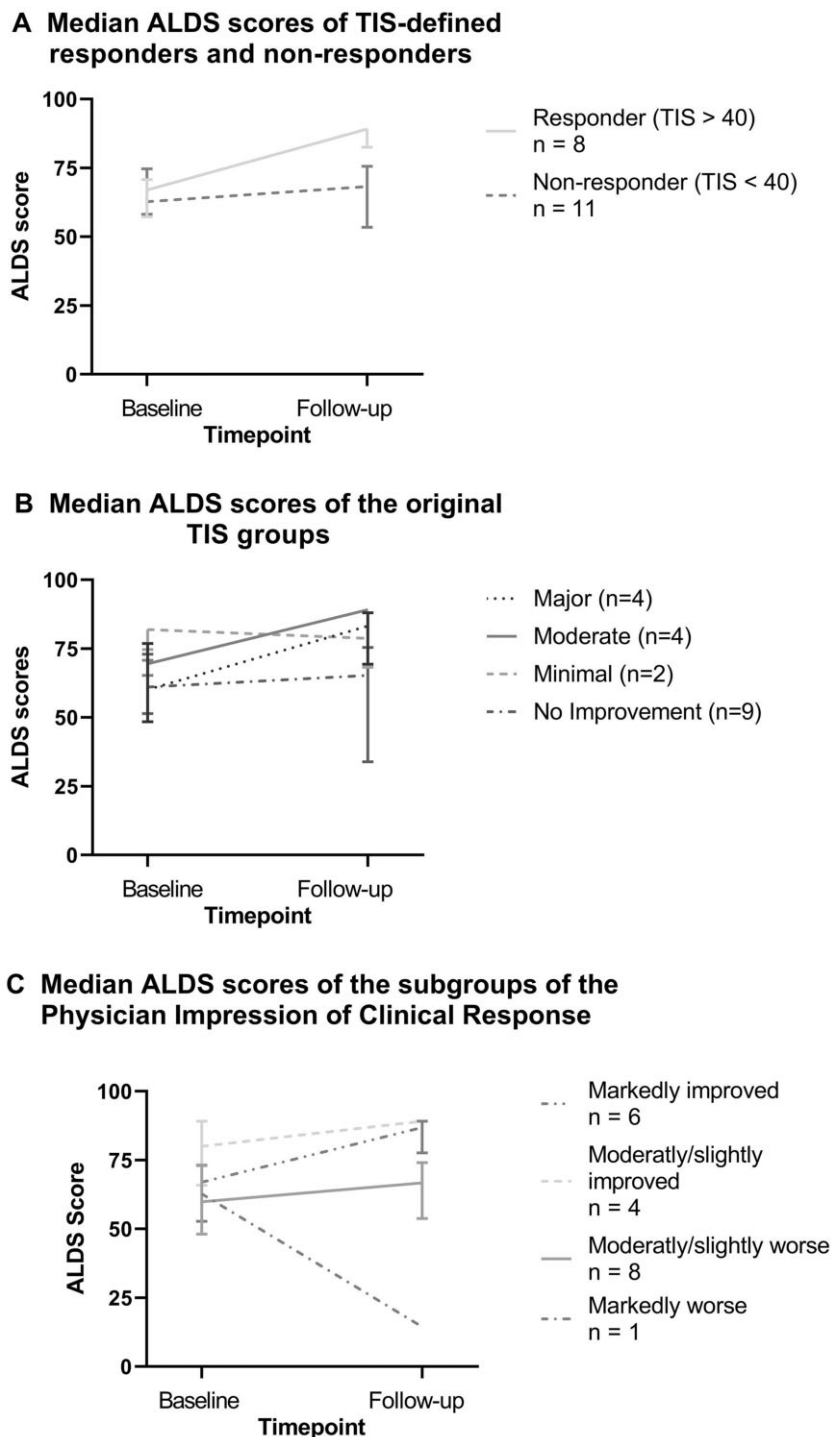
Thirdly, ALDS detects deterioration. This is relevant in clinical practice and clinical trials, both when a pharmaceutical compound is compared with standard of care and when dependency on maintenance treatment needs to be ascertained before an intervention [19].

Lastly, the ALDS does not require trained physicians to obtain the data as opposed to other patient reported outcome measures (PROMs), such as Myositis Activity Profile questionnaire [20], and can easily be filled in by the patient alone, e.g. as a digital survey.

Limitations of this study include a small sample size, heterogeneity among the myositis patients and absence of significant chronic disease damage (MRI data available, beyond the scope of this paper)—the latter may have contributed to the ceiling effect—and secondarily the use of a fixed-length ALDS. In future research, different sets of ALDS items for individual patients may be used, without losing the ability to compare scores [2, 21]. Further, ALDS could be transformed into a computerized adaptive test, with difficulty levels being automatically adapted, after each response [7, 14]. Future studies should include longer follow-up periods to further investigate responsiveness, and to include disability related to disease damage. As such, the ALDS, being a generic scale, may fulfil a need to measure disability, not only related to the myopathy, but also to the extramuscular disease manifestations (heart, lungs and joints).

In conclusion, our pilot study shows promising clinical properties of the ALDS in myositis, which warrants further investigation.

**Fig. 1** Known group validity of ALDS



(A) Median and IQR of ALDS scores of TIS-defined responders and non-responders. (B) Median and IQR of ALDS scores of original TIS groups. (C) Median and IQR of ALDS scores of the subgroups of the physician impression of clinical response. Data are median and IQR; follow-up = after nine weeks of IVIG monotherapy or premature ending of participation of the study. Note that data points in the graphs are based on median ALDS scores and numbers in the result section represent median(s) of individual change scores.



## Acknowledgements

We thank Ingrid de Groot for her help with selecting the ALDS items and Caroline Holtslag for her help with data entry.

All authors contributed to the conception or design of the work, or the acquisition, analysis, or interpretation of data. M.M., A.W.W., A.J.vdK. and J.R. wrote the initial draft drafting the work and all authors revised it critically for important intellectual content and gave final approval of this version to be published. All authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Several authors of this publication are members of the Netherlands Neuromuscular Center (NL-NMD) and the European Reference Network for rare neuromuscular diseases ERN-EURO-NMD.

**Collaborators:** Members of the Dutch Myositis Network collaborated on this project and gave final approval of this version to be published: Christiaan G.J. Saris, Department of Neurology, Radboud UMC, Donders Institute for Brain Cognition and Behaviour, Nijmegen, The Netherlands; Esther Brusse, Department of Neurology, Erasmus UMC, Rotterdam, The Netherlands; Jessica E. Hoogendijk, Department of Neurology, Brain Centre; Rudolf Magnus, UMC Utrecht, Utrecht, The Netherlands.

**Funding:** To carry out the work described in this article no specific funding was received from any bodies in the public, commercial or not-for-profit sectors. The IMMEDIATE study was funded by a grant from CSL Behring.

**Disclosure statement:** The authors have declared no conflicts of interest.

## Data availability statement

The data underlying this article will be shared on reasonable request to the corresponding author.

## Supplementary data

**Supplementary data** are available at *Rheumatology* online.

## References

- Rider LG, Aggarwal R, Machado PM *et al.* Update on outcome assessment in myositis. *Nat Rev Rheumatol* 2018;14:303–18.
- Benveniste O, Rider LG; ENMC Myositis Outcomes Study Group. 213th ENMC International Workshop: outcome measures and clinical trial readiness in idiopathic inflammatory myopathies, Heemskerk, The Netherlands, 18–20 September 2015. *Neuromuscul Disord* 2016;26:523–34.
- Aggarwal R, Rider LG, Ruperto N *et al.* 2016 American College of Rheumatology/European League Against Rheumatism criteria for minimal, moderate, and major clinical response in adult dermatomyositis and polymyositis: an international myositis assessment and clinical studies group/paediatric rheumatology international trials organisation collaborative initiative. *Arthritis Rheumatol* 2017;69:898–910.
- Agarwal SK, Monach PA, Docken WP, Coblyn JS. Characterization of relapses in adult idiopathic inflammatory myopathies. *Clin Rheumatol* 2006;25:476–81.
- Mecoli CA, Park JK, Alexanderson H *et al.* Perceptions of patients, caregivers, and healthcare providers of idiopathic inflammatory myopathies: an international OMERACT study. *J Rheumatol* 2019;46:106–11.
- Jabrayilov R, Emons WHM, Sijtsma K. Comparison of classical test theory and item response theory in individual change assessment. *Appl Psychol Meas* 2016;40:559–72.
- Holman R, Lindeboom R, Glas CAW, Vermeulen M, de Haan RJ. Constructing an item bank using item response theory: the AMC linear disability score project. *Health Services Outcomes Res Methodol* 2003;4:19–33.
- Holman R, Weisscher N, Glas CA *et al.* The Academic Medical Center Linear Disability Score (ALDS) item bank: item response theory analysis in a mixed patient population. *Health Qual Life Outcomes* 2005;3:83.
- Weisscher N, Post B, de Haan RJ *et al.* The AMC Linear Disability Score in patients with newly diagnosed Parkinson disease. *Neurology* 2007;69:2155–61.
- Weisscher N, Wijbrandts CA, De Haan R *et al.* The Academic Medical Center Linear Disability Score item bank: psychometric properties of a new generic disability measure in rheumatoid arthritis. *J Rheumatol* 2007;34:1222–8.
- Lim J, Eftimov F, Verhamme C *et al.* Intravenous immunoglobulins as first-line treatment in idiopathic inflammatory myopathies: a pilot study. *Rheumatology* 2021;60:1784–92.
- Herdman M, Gudex C, Lloyd A *et al.* Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36.
- Foundation. ER. EQ-5D-5L User Guide [updated 2019]. <https://euroqol.org/publications/user-guides>.
- Holman R, Glas CA, Lindeboom R, Zwiderman AH, de Haan RJ. Practical methods for dealing with ‘not applicable’ item responses in the AMC Linear Disability Score project. *Health Qual Life Outcomes* 2004;2:29.
- Akoglu H. User’s guide to correlation coefficients. *Turk J Emerg Med* 2018;18:91–3.
- Nunnally JC. *Psychometric theory*, 2nd edn. New York: Mcgraw Hill Book Company, 1978.
- Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology* 1990;1:43–6.
- Vanhoutte EK, Faber CG, Merkies IS; PeriNomS study group. 196th ENMC international workshop: outcome measures in inflammatory peripheral neuropathies 8–10 February 2013, Naarden, The Netherlands. *Neuromuscul Disord* 2013;23:924–33.
- van Schaik IN, Bril V, van Geloven N *et al.* Subcutaneous immunoglobulin for maintenance treatment in chronic

- inflammatory demyelinating polyneuropathy (PATH): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Neurol* 2018;17:35–46.
- 20 Tugwell P, Bombardier C, Buchanan WW *et al.* The MACTAR Patient Preference Disability Questionnaire—an individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. *J Rheumatol* 1987;14:446–51.
- 21 Hambleton RK, Jones RW. Comparison of classical test theory and item response theory and their applications to test development. *Educ Measur Issues Pract* 2005;12:38–47.