The Journal of
Rheumatology

Functional Index-3: A Valid and Reliable Functional Outcome Assessment Measure in Patients With Dermatomyositis and Polymyositis

Floranne C. Ernste[1] (ID), Christopher Chong[2] (ID), Cynthia S. Crowson[3] (ID), Tanaz A. Kermani[4] (ID), Orla Ni Mhuircheartaigh[5], and Helene Alexanderson[6] (ID)

ABSTRACT. Objective. Patients with dermatomyositis (DM) and polymyositis (PM) have reduced muscle endurance. The aim of this study was to streamline the Functional Index-2 (FI-2) by developing the Functional Index-3 (FI-3) and to evaluate its measurement properties, content and construct validity, and intra- and interrater reliability.

Methods. A dataset of the previously performed and validated FI-2 (n = 63) was analyzed for internal redundancy, floor, and ceiling effects. The content of the FI-2 was revised into the FI-3. Construct validity and intrarater reliability of FI-3 were tested on 43 DM and PM patients at 2 rheumatology centers. Interrater reliability was tested in 25 patients. The construct validity was compared with the Myositis Activities Profile (MAP), Health Assessment Questionnaire (HAQ), and Borg CR-10 using Spearman correlation coefficient.

Results. Spearman correlation coefficients of 63 patients performing FI-3 revealed moderate to high correlations between shoulder flexion and hip flexion tasks and similar correlations with MAP and HAQ scores; there were lower correlations for neck flexion task. All FI-3 tasks had very low to moderate correlations with the Borg scale. Intraclass correlation coefficients (ICC) of FI-3 tasks for intrarater reliability (n = 25) were moderate to good (0.88–0.98). ICC of FI-3 tasks for interrater reliability (n = 17) were fair to good (range 0.83–0.96).

Conclusion. The FI-3 is an efficient and valid method for clinically assessing muscle endurance in DM and PM patients. FI-3 construct validity is supported by the significant correlations between functional tasks and the MAP, HAQ, and Borg CR-10 scores.

Key Indexing Terms: idiopathic inflammatory myopathies, Functional Index-2, muscle endurance, outcome measures

[1]F.C. Ernste, MD, Division of Rheumatology, Department of Medicine, Mayo Clinic, Rochester, Minnesota, USA; [2]C. Chong, MD, Arthritis Associates and Osteoporosis Center of Colorado Springs, Colorado Springs, Colorado, USA; [3]C.S. Crowson, PhD, Department of Health Science Research, Division of Biostatistics, Mayo Clinic, Rochester, Minnesota, USA; [4]T.A. Kermani, MD, University of California at Los Angeles, California, USA; [5]O. Ni Mhuircheartaigh, MBBCh, Division of Rheumatology, Department of Medicine, Beacon Hospital, Dublin, Ireland; [6]H. Alexanderson, PhD, Department of Neurobiology, Care Science and Society, Division of Physiotherapy, and Department of Medicine, Division of Rheumatology Karolinska Institutet, and Function Area Occupational Therapy and Physical Therapy, Karolinska University Hospital, Solna, Stockholm, Sweden.

Address correspondence to Dr. F. Ernste, MD, 200 First Street SW, Rochester, MN 55905, USA. Email: ernste.floranne@mayo.edu.

Accepted for publication March 27, 2020.

The idiopathic inflammatory myopathies (IIM), dermatomyositis (DM) and polymyositis (PM), are inflammatory muscle diseases predominantly affecting the proximal skeletal muscles, causing muscle weakness, exercise intolerance, and functional disability[1,2,3]. Muscle inflammation leads to damage, scarring, and chronic muscle atrophy, with a lower proportion of type I muscle fibers responsible for muscle endurance relative to fast-twitch type II muscle fibers over time[4,5]. The clinical course is usually characterized by periods of remission and relapse; hence, clinical assessments of muscle strength and function must have the sensitivity to discriminate changes in disease activity to guide treatment decisions[6,7,8].

There are few valid and reliable tests available that measure muscle endurance as an assessment of disease activity and, by extension, physical function, that are representative of the ability to perform activities of daily living (ADL)[6,9]. The International Myositis Assessment and Clinical Studies Group (IMACS) recommends a 6-domain disease activity core set in clinical trials for the assessment of patients with IIM involving physician and patient global disease activity, muscle strength testing

by the manual muscle test (MMT), activity limitation by the Health Assessment Questionnaire (HAQ), muscle enzyme levels, and extraskeletal involvement and physical function[10,11]. The MMT is a widely used measure of muscle strength, but it does not reflect muscle endurance or correlate with the ability to perform ADL[6,8,9]. Indeed, patients with adult PM and DM are more limited in muscle endurance when assessed by the Functional Index 2 (FI-2) compared to muscle strength assessed by the MMT[8]; the FI-2 also seems to reflect self-reported physical function[12].

Tools to measure functional impairments and muscle endurance have been developed to clinically assess patients with IIM. The Adult Myopathy Assessment Tool (AMAT) holds promise as a 13-item performance test, designed to measure both physical function and endurance, that needs no specialized equipment to perform but requires 20–30 min to complete, and it has been validated with strong intrarater and interrater reliability scores[13]. The FI in myositis was the first outcome measure developed for assessing functional impairment in patients with DM and PM by testing 14 repetitive tasks of selected muscle groups in both upper and lower limbs; this test was effective in discriminating patients from healthy individuals[14]. However, the FI was time-consuming, with observed floor and ceiling effects in patients with mild to moderate impairment[15]. The ceiling effect occurs when patients of mild to moderate impairment cluster to the highest level of the measured outcome and therefore achieve the best score for the instrument. When a ceiling effect exists, any score beyond the upper limit cannot be measured; conversely, when an instrument has a floor effect, any score beyond the lowest limit cannot be measured[16]. Hence, the FI-2 was developed as a revision of the FI at the Karolinska University Hospital in Sweden, and it has been partially validated in patients with adult patients with DM and PM[17]. It involves testing 7 repetitive tasks performed either bilaterally or unilaterally with a metronome to standardize movement pace. The FI-2 is useful without ceiling or floor effects and is well tolerated by patients at varying stages of functional impairment. A prospective 7-week exercise study revealed the FI-2 to be sensitive to treatment outcomes with solid inter- and intrarater reliability[18]. The FI-2 demonstrates good to excellent interrater reliability (intraclass correlation coefficients [ICC] 0.86–0.99) and good construct validity, but it requires a maximum of 33 min to complete, and the concern for internal redundancy arises for some tasks such as shoulder abduction and step test[17,18]. The goal of revising the FI-2 to the FI-3 was to shorten the administration time of a functional assessment tool and to derive a total score (summation of the individual tasks divided by 3) of the instrument to make it more useful to clinicians to follow patient progress.

Therefore, we revised the FI-2 into the FI-3 as a clinical assessment tool to measure muscle endurance in patients with DM and PM for the purpose of incorporating into clinical practice and research trials in IIM. The objectives of our study were to validate the measurement properties of content and construct validity as well as intra- and interrater reliability of FI-3. Our prediction was there would be moderate to strong correlation of the FI-3 to measures of physical function and perceived exertion.

Our second hypothesis was that the FI-3 neck flexion would have the lowest correlation to the physical function measures, as the neck muscles may not be directly involved in the daily activities included in the Myositis Activities Profile (MAP) and the HAQ.

MATERIALS AND METHODS

The study was approved by the Mayo Clinic institutional review and ethics board at Mayo Clinic Rochester, Minnesota (12-007485), and the local ethics committee at Karolinska Institutet in Stockholm, Sweden. All subjects gave informed written consent prior to participation.

Study design. This was a cross-sectional study of a cohort of patients with DM or PM at varying stages of disease (active or in remission). They were recruited from the rheumatology clinics at the Karolinska University Hospital and Mayo Clinic Rochester.

Inclusion and exclusion criteria. Inclusion criteria were patients aged 18 years or over fulfilling Bohan and Peter criteria for DM or PM (at least 3 of 5 criteria needed for a probable diagnosis)[1]. Exclusion criteria were a diagnosis of inclusion body myositis, juvenile DM, severe pulmonary hypertension, acute fractures, or severe osteoporosis.

Patients. Three cohorts of patients were included in this study. Cohort 1 represented all patients performing the FI-2 during their early follow-up at the Karolinska University Hospital during 2006 (n = 63; site 1). Data were retrieved from the Swedish Myositis Register in 2010. Cohort 2 represented patients recruited at a second institution, Mayo Clinic Rochester, from 2010 to 2012, performing the FI-3 at site 2. Cohort 3 were patients performing FI-3 at site 1. These patients were seen on the same day as part of their scheduled follow-up visits.

Methods. The study was performed using the original FI-2 as the foundation for revision into the FI-3. The original FI-2 involves repetitive movements in 7 muscle groups to determine muscle endurance. It is made up of the following 7 tasks testing dynamic repetitive muscle function: shoulder flexion with 1-kg weight cuff on wrists, shoulder abduction, neck flexion/ head lift, hip flexion, step test, heel lift, and toe lift. The patient performed many repetitions possible to a maximum of 3 min per task (60–120 repetitions) with a metronome (40 beats/min generating 20 repetitions) to standardize the pace. A maximum of 3 min per task was allotted to reach the maximal number of repetitions. Patients did 5 learning repetitions to enhance performance. The entire test was performed in 1 sitting.

The study investigators met to scrutinize the goals of the tasks of the FI-2 *a priori* and then performed a separate analysis of tasks for internal redundancy. In an effort to avoid bias, we sought advice on the development of the FI-3 through informal discussions among the authors of the FI-2, relating to goals of the FI-2. Tasks that were felt to be essential to the functional assessment of DM/PM patients were the forward flexion, neck flexion, and hip flexion tasks[19]. The step test task was removed due to the inclusion of nonessential elements of cardiovascular stress, balance, and coordination. The heel lift and toe lift tasks were also removed, as they test distal muscle function, which may not be limited in patients with PM and DM. The total score (summation of the individual tasks divided by 3) was developed for ease of use. Table 1 illustrates the differences between the FI-2 and the FI-3.

For determining the FI-3 validation, study investigators at both sites participated in a 1-h training session discussing the process of evaluating patients using the revised FI-2 and the scoring method prior to beginning the data collection (Supplementary Data 1 and 2, available with the online version of this article). Scoring was based on the number of correctly performed repetitions with a score varying from 0 to 60. A score of 60 in a task reflects normal muscle endurance. The total score of the FI-3 was calculated on 1 side as right shoulder flexion, right hip flexion, and neck flexion, and divided by 3.

For assessing the interrater reliability, a total of 17 subjects performed the FI-3 twice at the same visit with a resting span of 30–60 min. The test was performed, randomly led by assessor 1 and assessor 2 at each site. In

Table 1. Characteristics of FI-2 and FI-3.

	FI-2	FI-3
Total no. tasks	7	3
Performed unilaterally or bilaterally	Yes	Yes
Tasks involved	Shoulder flexion, shoulder abduction, neck flexion, hip flexion, step test, heel lift, toe lift	Shoulder flexion, neck flexion, hip flexion
Pace standardized with metronome (40 or 80 beats/min)	Yes	Yes
Time per task, min	3	3
Total time required to perform all tasks unilaterally/ bilaterally, min	21/33	9/15
Repetitions per task	60–120	60
Total score derived	No	Yes

FI: Functional Index.

the first session, subjects performed bilateral shoulder flexion (sequentially), neck/flexion/head lift, and bilateral hip flexion (sequentially). The Borg CR-10 scale was filled out after each task. The Borg CR-10 is a category scale to rate perceived muscular exertion ranging from 0 = nothing at all to 10 = very strong[20]. After completing the first session, patients filled out the MAP[21,22] and the HAQ[23,24].

The MAP is a myositis-specific questionnaire developed and validated in Sweden measuring limitations in daily life activities. Thirty-one items are scored on a 7-point Likert scale from 1 to 7: 1 = no difficulty to perform and 7 = impossible to perform. The MAP has also been validated in patients with DM/PM in the United States[21,22]. The HAQ is an arthritis-specific 20-item questionnaire assessing functional ability and it is also recommended for use in myositis, including inflammatory myopathies, following the effect of exercise interventions[23,24].

For assessing the intrarater reliability, a total of 29 subjects (cohort 2) at site 2 and 14 patients (cohort 3) at site 1 performed the FI-3 twice, with a resting interval of 30–60 min by assessor 1 at site 2. At site 1, the FI-3 was performed twice within 4–7 days by the same assessor. After completing each task, perceived exertion was again rated using the Borg CR-10 scale.

Statistical analysis. Descriptive data are presented using percentages, means, medians, SD, and ranges, with graphic depiction using box plots. Spearman correlation coefficient was used to assess internal redundancy (how all tasks correlate to each other) and internal consistency (how each task correlates to a subscore of upper or lower extremities) of the FI-2. Correlation coefficients Rs > 0.90 were considered to indicate redundancy and Rs < 0.60 to indicate poor consistency. In analysis for construct validity, correlation coefficients of Rs 0–0.25 were considered as no or very low, Rs 0.25–0.49 as low, Rs 0.50–0.69 as moderate, Rs 0.70–0.89 as high, and Rs 0.90–10.00 as very high[25]. ICC were calculated for intrarater and interrater reliability. ICC < 0.75 indicate low to fair reliability, and those > 0.75 indicate good to excellent reliability. The reliability of the total score and error of measurement in all tasks were calculated. The level of significance was accepted as *P* < 0.05 for all comparisons. All data were analyzed using SAS version 9.4 (SAS Institute Inc.) and R version 3.2.3 (R Foundation for Statistical Computing).

RESULTS

Content validity. The FI-2 tasks were scrutinized, and the range of FI-2 mean scores was 8.2 ± 10.8 to 40.2 ± 29.5. No ceiling effects (defined as median values of 20–50 percentile of total variation of values) were observed in any tasks. Correlation coefficients for analysis of internal redundancy for FI-2 tasks in the upper extremities varied between Rs 0.89–0.94, with the highest correlations between shoulder flexion and shoulder abduction. The shoulder flexion task ranged between 0–60 with no tendency toward ceiling effect, while the median shoulder abduction equaled the maximal number of repetitions at 60; therefore, the shoulder abduction latter task was removed due to internal redundancy. Coefficients for analysis of tasks in the lower extremities varied between Rs 0.63–0.84. The cutoff for exclusion due to poor internal redundancy was Rs < 0.60 (Figure 1).

Clinical characteristics of patients in cohorts 2 and 3. Twenty-nine patients comprised cohort 2; 14 patients made up cohort 3. Table 2 describes the clinical characteristics of these patients. In cohort 2, the mean age was 57.9 (SD 11.2) years and 66% were female. In cohort 3, the mean age was 62.0 (SD 13.5) years and 71% were female. In cohort 2, 52% had DM, 21% had PM, and 28% had antisynthetase syndrome. Cohort 3 included 21% DM and 79% PM patients. The median creatine kinase level was 102 U/L (range 43–850 U/L) in cohort 2 and 97 (range 47–1470) U/L in cohort 3. The median myositis disease duration was 5 years in cohort 2 and 11.5 years in cohort 3. The median MAP score in session 1 was 2 (range 1–5). The median HAQ score in session 1 was 0.6 [range 0–2; cohort 2 = 0.8 (range 0–2) and cohort 3 = 0.6 (range 0–2)]. Other than disease duration, the patient characteristics were similar in both cohorts, so the cohorts were combined for further analyses.

Construct validity. To measure construct validity, the correlations between FI-3 tasks (% of maximum no. repetitions) and MAP scores (n = 42), HAQ scores (n = 39), and Borg CR-10 scores (n = 43) in session 1 were assessed (Table 3). The correlations between physical function and FI-3 total score were the following: MAP (–0.67, *P* < 0.001), HAQ (–0.72, *P* < 0.001). There were moderate to high correlations between shoulder flexion and hip flexion tasks and the MAP and the HAQ, with lower correlations for the neck flexion task. All FI-3 tasks had very low to moderate correlations with the Borg scale.

Intrarater reliability. Intrarater reliability of the FI-3 tasks (% maximum no. repetitions) and Borg CR-10 scores was assessed in 25 subjects in cohorts 2 and 3, who were observed twice by the same rater (Table 4).

For FI-3, the ICC was 0.96 (95% CI 0.92–0.98) for the right shoulder flexion task and 0.92 (95% CI 0.83–0.97) for the left shoulder flexion task. The ICC was 0.94 (95% CI 0.87–0.97) for the neck flexion task. The ICC was 0.88 (95% CI 0.74–0.94) for the right hip flexion task and 0.93 (95% CI 0.86–0.97) for the left hip flexion task. The measurement of error for each FI-3 task was calculated: right shoulder flexion (6.8), left shoulder flexion (9.3), neck flexion (9.0), right hip flexion (13.2), and left hip flexion (10.0). The ICC for the total score was 0.95 (95% CI 0.89–0.98).

Interrater reliability. To determine interrater reliability, 17 subjects in cohorts 2 and 3 were observed by 2 different assessors
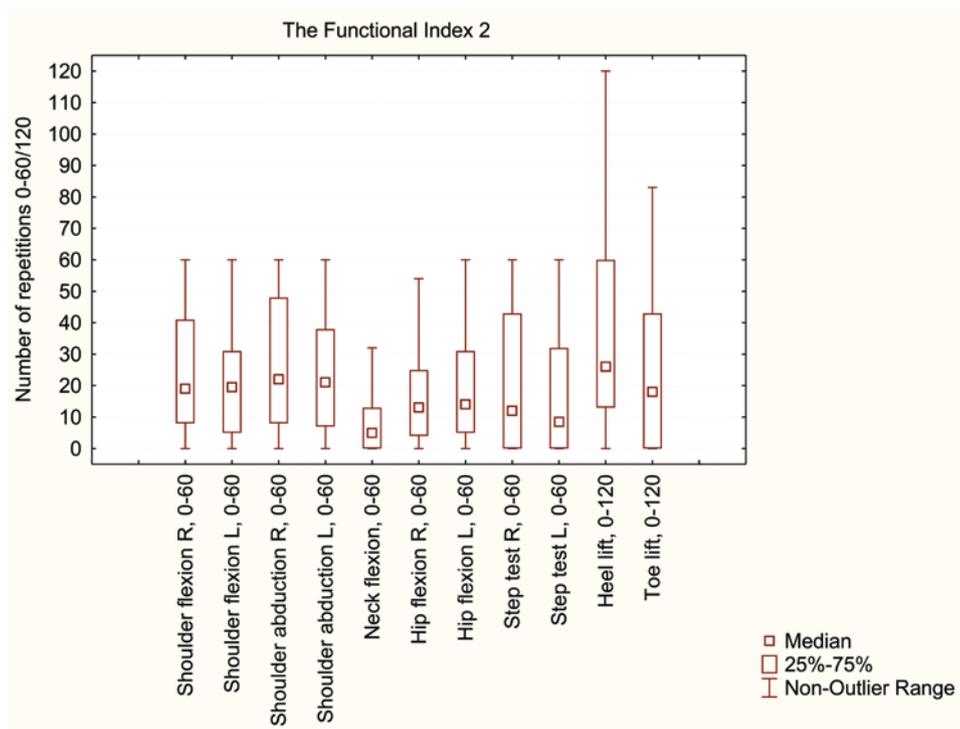
Figure 1. Descriptive data of 63 subjects who participated in the FI-2 tests. Maximal FI-2 score is 60 for all tasks, except the heel lift and toe lift tasks (120 is maximal). FI-2: Functional Index 2.

Table 2. Demographic data of 29 patients in cohort 2 and 14 patients in cohort 3.

Clinical Characteristics	Cohort 2, n = 29	Cohort 3, n = 14
Age, yrs, median (range)	60 (31–76)	61.5 (43–86)
Mean (SD)	57.9 (11.2)	62.0 (13.5)
Myositis duration, yrs, median (range)	5 (1–22)	11.5 (5–43)
Mean (SD)	6.1 (5.3)	15.4 (10.9)
Creatine kinase, U/L, median (range)	102 (43–850)	97 (47–1470)
Sex, female	19 (66)	10 (71)
Diagnosis		
PM	6 (21)	11 (79)
DM	15 (52)	3 (21)
Antisynthetase syndrome	8 (28)	0 (0)
MAP score, median (range)	2 (1–5)	2 (1–5)
HAQ score, median (range)	0.8 (0–2)	0.6 (0–2)

Values are n (%) unless otherwise specified. DM: dermatomyositis; HAQ: Health Assessment Questionnaire; MAP: Myositis Activities Profile; PM: polymyositis.

in session 1 and session 2. Table 5 shows the medians and ranges for each session and the ICC of the assessments. The measurement of error for each FI-3 task was calculated: right shoulder flexion (15.5), left shoulder flexion (13.9), neck flexion (14.1), right hip flexion (8.4), and left hip flexion (7.4). The ICC for the total score was 0.93 (95% CI 0.83–0.98).

DISCUSSION
The FI-3 represents a streamlined version of the FI-2 in patients with DM and PM taking a maximum of 15 min with bilateral assessment and a maximum of 9 min with unilateral assessment. The FI-3 is a reliable and valid method for the functional assessment of DM/PM patients for muscle impairment of the major muscle groups in the neck and upper and lower extremities in patients at various stages of disease. The ICC scores for inter-/intrarater reliability for the FI-3 tasks were good to excellent; there were significant moderate to high associations with the MAP and HAQ scores to support construct validity. The neck flexion, bilateral shoulder flexion, and bilateral hip flexion tasks are in line with the common disease phenotype of bilateral proximal muscle involvement in patients with DM/PM[2]. All tasks of the FI-3 had good to excellent intra- and interrater reliability. As expected, the interrater ICC values were similar, but slightly lower than those for intrarater reliability. This indicates that the FI-3 is easy to perform, as it does not require more than 1 training session between assessors. There were no reported adverse effects from use of FI-3 in our patients. Further, the FI-3 does not require specialized equipment or sophisticated training to conduct.

The FI-3 is to some extent similar to the AMAT[13]; however, we anticipate that the advantage of the FI-3 over the AMAT is that it may be easier to use due to fewer tasks involved, and therefore less training is needed. Similar to the FI-3, the AMAT features functional and endurance tasks, yet the original item development of the AMAT involved adults with DM/PM and inclusion body myositis, and included the arm raise, modified push-up, sit to stand, sit-up, step-up, supine to prone, supine to sit, head elevation, repeated heel rise, hip flexion, and knee extension. While all are important tasks to assess function and

Table 3. Spearman correlations between MAP and HAQ scores and Borg CR-10 scores with FI-3 task scores (% of maximum number of repetitions) in session 1 for 43 patients in cohorts 2 and 3.

FI-3 Task Score	MAP Score, n = 42		HAQ Score, n = 39		Borg CR-10 Scores, n = 43	
	Correlation	P	Correlation	P	Correlation	P
Shoulder flexion right	−0.62	< 0.001	−0.68	< 0.001	−0.48	0.001
Shoulder flexion left	−0.58	< 0.001	−0.60	< 0.001	−0.32	0.039
Neck flexion	−0.38	0.011	−0.43	0.006	−0.30	0.049
Hip flexion right	−0.70	< 0.001	−0.76	< 0.001	−0.21	0.16
Hip flexion left	−0.61	< 0.001	−0.68	< 0.001	−0.10	0.52
Total score	−0.67	< 0.001	−0.72	< 0.001	NA	NA

FI-3: Functional Index 3; HAQ: Health Assessment Questionnaire; MAP: Myositis Activities Profile; NA: not available.

Table 4. Intrarater reliability of FI-3 task score (% of maximum number of repetitions) and Borg CR-10 scores for 25 patients in cohorts 2 and 3.

	Session 1 Median (Range)	Session 2 Median (Range)	Difference Median (Range)	ICC (95% CI)	Measurement Error
FI-3 Tasks					
Shoulder flexion right	100 (0–100)	100 (5–100)	0 (−28 to 25)	0.96 (0.92–0.98)	6.8
Shoulder flexion left	100 (8–100)	67 (7–100)	0 (−43 to 11)	0.92 (0.83–0.97)	9.3
Neck flexion	40 (0–100)	35 (0–100)	0 (−38 to 20)	0.94 (0.87–0.97)	9.0
Hip flexion right	50 (0–100)	33 (0–100)	0 (−57 to 33)	0.88 (0.74–0.94)	13.2
Hip flexion left	45 (0–100)	38 (0–100)	0 (−48 to 40)	0.93 (0.86–0.97)	10.0
Total score	69 (0–100)	62 (3–100)	−1 (−32 to 15)	0.95 (0.89–0.98)	6.8
Borg CR-10 scores					
Shoulder flexion right	3 (0.5–9)	4 (0.5–9)	0 (−2 to 4)	0.76 (0.54–0.89)	1.0
Shoulder flexion left	4 (0.5–9)	4 (0.5–9)	0 (−2 to 4)	0.85 (0.68–0.93)	0.9
Neck flexion	5 (3–10)	5 (3–9)	0 (−4 to 5)	0.53 (0.17–0.76)	1.6
Hip flexion right	6 (3–10)	6 (0.5–10)	0 (−6 to 2)	0.72 (0.46–0.86)	1.2
Hip flexion left	6 (4–10)	7 (2–10)	0 (−6 to 3)	0.68 (0.38–0.84)	1.2

FI-3: Functional Index 3; ICC: intraclass correlation coefficient.

Table 5. Interrater reliability of FI-3 task score (% of maximum number of repetitions) and Borg CR-10 scores for 17 patients in cohorts 2 and 3.

	Session 1 Median (Range)	Session 2 Median (Range)	Difference Median (Range)	ICC (95% CI)	Measurement Error
FI-3 Tasks					
Shoulder flexion right	83 (0–100)	72 (3–100)	0 (−73 to 24)	0.83 (0.59–0.93)	15.5
Shoulder flexion left	78 (8–100)	60 (7–100)	0 (−65 to 9)	0.83 (0.59–0.94)	13.9
Neck flexion	42 (0–100)	38 (0–100)	0 (−19 to 55)	0.85 (0.63–0.94)	14.1
Hip flexion right	32 (0–100)	33 (2–100)	0 (−40 to 12)	0.95 (0.88–0.98)	8.4
Hip flexion left	35 (0–100)	33 (0–100)	0 (−27 to 14)	0.96 (0.90–0.99)	7.4
Total score	51 (0–100)	47 (2–100)	0 (−28 to 20)	0.93 (0.83–0.98)	8.8
Borg CR-10 scores					
Shoulder flexion right	4 (0–7)	4 (0–10)	0 (−2 to 4)	0.78 (0.50–0.92)	1.2
Shoulder flexion left	4 (0–9)	4 (0–10)	0 (−1 to 4)	0.89 (0.74–0.96)	0.9
Neck flexion	5 (0–9)	6 (0–10)	1 (−1 to 7)	0.64 (0.24–0.86)	1.5
Hip flexion right	6 (0–10)	5 (0–10)	0 (−7 to 2)	0.75 (0.44–0.90)	1.4
Hip flexion left	5 (0–10)	5 (0–10)	0 (−4 to 2)	0.88 (0.69–0.95)	1.0

FI-3: Functional Index 3; ICC: intraclass correlation coefficient.

endurance, they are not specific for patients with DM/PM; they also include tasks for testing distal muscle strength and coordination that are relevant to patients with myopathies with distal muscle involvement. Moreover, we speculate that due to potential patient safety concerns, healthcare providers may be hesitant to administer certain tasks that require balance and coordination skills to patients with significant disabilities. Finally, the advantage of the FI-3 is that if time constraints exist, a single task can

be measured to determine a baseline, and then performance of that task may be followed over time for discrimination of change since each task has been validated separately.

Given the number of current tools (i.e., AMAT, FI-2, and FI-3) available to assess the functional status and endurance of patients with myositis, there is a potential concern that there are many overlapping similarities. Hence, it could be challenging to decide which tool to use for clinical and/or research purposes. We believe that all the tools retain unique features that address different endurance tasks: The AMAT and FI-2 represent a comprehensive assessment of multiple muscle groups while incorporating balance and coordination skills, and the FI-3 offers a streamlined version of 3 major tasks with a derivation of a total score for ease of assessment of patient progress.

There are several potential limitations to this study. First, in session 2, the smaller sample size in the Mayo Clinic cohort may have affected the reliability scores. FI-3 task scores in the site 2 cohort were more likely to decrease from session 1 to session 2, whereas at site 1, the FI-3 scores were more likely to increase between the 2 sessions. This may be due to fatigue, since the site 2 cohort performed the 2 sessions on the same day whereas the site 1 cohort performed the tasks on 2 separate days. There may not have been adequate time to rest between sessions (30–45 min, max). Second, there may be a ceiling effect (median values equal 20–50 percentile of the total variation of values) in the shoulder flexion task in patients at both sites. It was unexpected that both groups would perform so well in the shoulder flexion task. We expected that the most demanding tasks were hip flexion and neck flexion[19]. Moreover, we observed that FI-3 tasks had low to moderate correlations to the Borg CR10 scale. Initially, the low correlations with the Borg index were surprising to us, but in reflection seem plausible, since measures of perceived exertion are difficult to quantify given the subjectivity of patients self-rating their difficulties in performance of tasks, possibly related to motivational factors and attitudes. We speculate that patients may not have been motivated to perform as anticipated, or they perceived tasks to be more strenuous than expected. Some patients may experience a very fast depletion of muscle function, from one repetition to the next; in these cases, patients do not experience high muscle exertion even though they are unable to continue performing additional repetitions. This could also explain the low correlations between number of performed repetitions and the perceived exertion. Potentially, the construct validity of our study could be compromised with choosing the Borg scale to assess for correlations of tasks. However, we believe that the Borg scale will help the assessor to understand why the patient is unable to continue the test, for reasons such as fatigue, pain, or low motivation. A low perceived exertion warrants further questions to assess for limiting factors such as cardiovascular, pulmonary, or musculoskeletal conditions that may lead to termination of testing. Finally, patients filled out the MAP and HAQ forms following completion of the FI-3; hence, their performances on the FI-3 may have influenced their self-assessments, including their MAP and HAQ scores.

In conclusion, the FI-3 is a valid and reliable tool to safely assess muscle endurance in patients with PM/DM. We anticipate the FI-3 can be incorporated into the routine functional assessment reflecting stamina and muscle endurance. We suggest that the FI-3 tasks be used on the dominant side as it requires only 10 min to perform and scores can be calculated separately for each task or as a total score. The FI-3 can complement the FI-2 and the AMAT, as well as the proposed IMACS core set of outcome measures[13,17].

REFERENCES

1. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). N Engl J Med 1975;292:34 4-7.
2. Harris-Love MO, Shrader JA, Koziol D, Pahlajani N, Jain M, Smith M, et al. Distribution and severity of weakness among patients with polymyositis, dermatomyositis and juvenile dermatomyositis. Rheumatology 2009;48:134-9.
3. Wiesinger GF, Quittan M, Nuhr M, Volc-Platzer B, Ebenbichler G, Zehetgruber M, et al. Aerobic capacity in adult dermatomyositis/ polymyositis patients and healthy controls. Arch Phys Med Rehabil 2000;81:1-5.
4. Dastmalchi M, Alexanderson H, Loell I, Ståhlberg M, Borg K, Lundberg IE, et al. Effect of physical training of the proportion of slow-twitch type I muscle fibers, a novel nonimmune-mediated mechanism for muscle impairment in polymyositis or dermatomyositis. Arthritis Rheum 2007;57:1303-10.
5. Loell I, Helmers SB, Dastmalchi M, Alexanderson H, Munters LA, Nennesmo I, et al. Higher proportion of fast-twitch (type II) muscle fibres in idiopathic inflammatory myopathies — evident in chronic but not in untreated newly diagnosed patients. Clin Physiol Funct Imaging 2011;31:18-25.
6. Agarwal S, Kiely PD. Two simple, reliable and valid tests of proximal muscle function, and their application to the management of idiopathic inflammatory myopathies. Rheumatology 2006; 45:874-9.
7. Miller FW. New approaches to the assessment and treatment of the idiopathic inflammatory myopathies. Ann Rheum Dis 2012;71 Suppl 2:i82-5.
8. Rider LG, Giannini EH, Harris-Love M, Joe G, Isenberg D, Pilkington C, et al. Defining clinical improvement in adult and juvenile myositis. J Rheumatol 2003;30:603-17.
9. Harris-Love MO. Physical activity and disablement in the idiopathic inflammatory myopathies. Curr Opin Rheumatol 2003;15:679-90.
10. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-45.
11. Miller FW, Rider LG, Chung YL, Cooper R, Danko K, Farewell V, et al; International Myositis Outcome Assessment Collaborative Study Group. Proposed preliminary core set measures for disease outcome assessment in adult and juvenile idiopathic inflammatory myopathies. Rheumatology 2001;40:1262-73.
12. Alexanderson H, Regardt M, Ottoson C, Alemo Munters L, Dastmalchi M, Dani L, et al. Muscle strength and muscle endurance during the first year of treatment of polymyositis and dermatomyositis: a prospective study. J Rheumatol 2018; 45:4538-46.
13. Harris-Love MO, Joe G, Davenport TE, Koziol D, Rose KA, Shrader JA, et al. Reliability of the adult myopathy assessment tool in individuals with myositis. Arthritis Care Res 2015;67:563-70.
14. Josefson A, Romanus E, Carlsson J. A functional index in myositis. J Rheumatol 1996;23:1380-4.
15. Alexanderson H, Stenström CH, Jenner G, Lundberg I. The safety of a resistive home exercise program in patients with recent onset active polymyositis or dermatomyositis. Scand J Rheumatol 2000;29:295-301.

16. Everitt BS and Skrondal A. The Cambridge Dictionary of Statistics. 4th edition. Cambridge, UK: Cambridge University Press; 2010;84.

17. Alexanderson H, Broman L, Tollbäck A, Josefson A, Lundberg I, Stenström CH. Functional index-2: validity and reliability of a disease-specific measure of impairment in patients with polymyositis and dermatomyositis. Arthritis Rheum 2006;55:114-22.

18. Alexanderson H, Dastmalchi M, Esbjörnsson-Liljedahl M, Opava CH, Lundberg IE. Benefits of intensive resistance training in patients with chronic polymyositis or dermatomyositis. Arthritis Rheum 2007;57:768-77.

19. Harris-Love MO, Shrader JA, Koziol D, Pahlajani N, Jain M, Smith M, et al. Distribution and severity of weakness among patients with polymyositis, dermatomyositis and juvenile dermatomyositis. Rheumatology 2009;48:134-9.

20. Borg GA. Psychophysical bases of perceived exhaustion. Med Sci Sports Exerc 1982;14:377-81.

21. Alexanderson H, Lundberg IE, Stenström CH. Development of the Myositis Activities Profile-validity and reliability of a self-administered questionnaire to assess activity limitations in patients with polymyositis/dermatomyositis. J Rheumatol 2002:29:2386-92.

22. Alexanderson H, Reed AM, Ytterberg SR. The Myositis Activities Profile-initial validation for assessment of polymyositis/dermatomyositis in the USA. J Rheumatol 2012;39:2134-41.

23. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales. J Rheumatol 1982:789-93.

24. Ramey DR, Raynauld JP, Fries JF. The Health Assessment Questionnaire 1992: status and review. Arthritis Care Res 1992:119-29.

25. Munro BH. Statistical Methods for Health Care research. 3rd ed. Philadelphia: Lippincott; 1997;235.