Seminars in Arthritis and Rheumatism 68 (2024) 152534

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit

Construct validity of PROMIS pain interference, fatigue, and physical function as patient-reported outcomes in adults with idiopathic inflammatory myopathies: An international study from the OMERACT myositis working group

Ellen Romich [a,1,*], Didem Saygin [b,1], Dana DiRenzo [a], Christopher A. Mecoli [c], Ingrid de Groot [d], Karin Lodin [e], Malin Regardt [f], Catherine Sarver [g], Ju Yeon Kim [h], Jin Kyun Park [h], Kelly Beer [i], Merrilee Needham [i], Helene Alexanderson [j], Lisa Christopher-Stine [c], Marianne de Visser [k,2], Joost Raaphorst [k,2], on behalf of the OMERACT Myositis Working Group

[a] Division of Rheumatology, Department of Medicine, University of Pennsylvania, USA
[b] Division of Rheumatology, Department of Medicine, University of Pittsburgh, USA
[c] Division of Rheumatology, Department of Medicine, Johns Hopkins University, USA
[d] Patient Research Partner, the Netherlands
[e] Division of Rheumatology, Department of Medicine, Solna, Karolinska Institutet, Karolinska University, Stockholm, Sweden
[f] Department of Neurobiology, Care Sciences and Society, Division of Occupational Therapy, Karolinska Institutet and Medical Unit Occupational Therapy and Physical Therapy, Karolinska University Hospital, Stockholm, Sweden
[g] Patient Research Partner, USA
[h] Division of Rheumatology, Department of Internal Medicine, Seoul National University Hospital and College of Medicine, Seoul, South Korea
[i] Department of Neurology, Fiona Stanley Hospital, IIID Murdoch University and University of Notre Dame, Perth, Australia
[j] Medical Unit Occupational Therapy and Physical Therapy, Karolinska University Hospital and Division of Rheumatology, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden
[k] Department of Neurology, Amsterdam UMC, University of Amsterdam, Amsterdam Neuroscience, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Myositis
Patient-reported outcome
Outcome measurement
Quality of life Measure
OMERACT

ABSTRACT

Background: Validated patient-reported outcome measures to assess disease impact in patients with adult idiopathic inflammatory myopathies (IIMs) are needed. The objective of this study was to assess the construct validity of PROMIS Pain Interference, Fatigue, and Physical Function measures in comparison with core disease activity measures.

Methods: Adults with IIM, excluding inclusion body myositis, from OMERACT Myositis Working Group (MWG) clinic sites completed PROMIS Short Form v1.0—Pain Interference 6a, PROMIS Short Form v1.0—Fatigue 7a, and PROMIS Short Form v2.0—Physical Function 8b measures. Core disease activity measures including patient and physician global disease activity assessments, manual muscle testing, serum creatine kinase activity, and Health Assessment Questionnaire Disability Index (HAQ-DI) were simultaneously assessed. To evaluate construct validity, a priori hypotheses for the expected correlations between PROMIS measures, age, and core disease measures were determined by >70 % agreement among MWG members and were compared against observed Pearson's correlations. Internal consistency of items and floor or ceiling effects for the PROMIS measures were also assessed. Subgroup analysis according to IIM subtype (dermatomyositis vs. non-dermatomyositis IIM) was performed.

Results: 135 adults with IIM from 5 countries across North America, Europe, Asia, and Australia were included. For construct validity, a priori hypotheses were confirmed for 5 of 6 (83 %) PROMIS Pain Interference, 4 of 5 (80 %) PROMIS Fatigue, and 3 of 4 (75 %) PROMIS Physical Function correlations. Internal consistency was high for each PROMIS measure (Cronbach's alpha >0.9). Ceiling effects were observed only for PROMIS Pain

* Corresponding author at: University of Pennsylvania, Department of Rheumatology, White Building, 5th Floor, 3400 Spruce Street, Philadelphia, PA 19104, USA.
E-mail address: ellen.romich@pennmedicine.upenn.edu (E. Romich).
[1] Co-first authors.
[2] Co-senior authors

https://doi.org/10.1016/j.semarthrit.2024.152534

Available online 10 August 2024

Interference, with low/no pain in 29 % of patients. Subgroup analysis between dermatomyositis ($n = 65$) and non-dermatomyositis ($n = 70$) subtypes demonstrated similar correlations between PROMIS measures and disease activity measures.
Conclusions: PROMIS Short Form v1.0—Pain Interference 6a, PROMIS Short Form v1.0—Fatigue 7a, and PROMIS Short Form v2.0—Physical Function 8b measures demonstrate strong construct validity when compared to core disease activity measures in IIM, with consistent results across IIM subtypes. These findings support the use of these selected PROMIS measures to assess core domains of interest for measuring life impact in IIMs.

Introduction

Idiopathic inflammatory myopathies (IIMs) are a group of systemic autoimmune diseases including dermatomyositis, immune-mediated necrotizing myopathy, anti-synthetase syndrome, polymyositis, and overlap myositis. While muscle weakness is a common presenting feature in IIMs, extra-muscular manifestations including interstitial lung disease, skin involvement, and arthritis are frequently observed [1]. Despite treatment, many patients develop physical limitations and reduced quality of life [2,3], necessitating the need for validated outcome measures to assess aspects of myositis, including patient-reported outcomes (PROs) that capture the patient experience and disease impact.

Outcome Measures in Rheumatology (OMERACT) is a global volunteer-driven, non-profit organization that brings together researchers, clinicians, and patients to facilitate outcome measure development and validation using a rigorous framework [4]. The OMERACT Myositis Working Group, established in 2011, is an international community that includes clinicians (rheumatologists, neurologists, physical and occupational therapists), methodologists, and patient research partners from Australia, Canada, the Netherlands, South Korea, Sweden, and the United States (US) [5]. Over the past decade, efforts from this working group through focus groups and modified Delphi exercises following OMERACT methodology led to the identification of pain interference, fatigue, and physical function as core domains of interest to assess life impact of IIMs, and these domains should be prioritized for assessment in clinical trials [6–9]. Subsequent efforts identified three Patient-Reported Outcomes Measurement Information System (PROMIS) measures as representative PRO instruments for these core domains: PROMIS Short Form v1.0—Pain Interference 6a, PROMIS Short Form v1.0—Fatigue 7a, and PROMIS Short Form v2.0—Physical Function 8b [10]. These measures have strong test-retest reliability, internal consistency, and responsiveness [11,12]. Meaningful thresholds for change have also been demonstrated [12]. Construct validity, or the degree to which the scores of these PROMIS measures relate to other measures in a manner consistent with a priori expectations [13], was examined in an international, though predominantly US-based, cohort of adults with IIM with comparison to quality of life measures including measures of sleep, anxiety, pain intensity, physical activity, and disability, and results demonstrated strong construct validity [11].

Using a separate cohort of adults with IIM with larger international representation, this study aims to further validate the internal consistency and construct validity of PROMIS Pain Interference, Fatigue, and Physical Function measures in comparison with the International Myositis Assessment and Clinical Studies (IMACS) core set measures used to assess disease activity in IIMs. The core set measures include patient and physician global assessments, extra-muscular global disease activity, manual muscle strength testing, serum muscle enzyme activity, and Health Assessment Questionnaire Disability Index (HAQ-DI) [14]. We also aimed to evaluate whether the construct validity of the three representative PROs differs by IIM subtype and to compare two instruments, the PROMIS Physical Function Short Form 8b and HAQ-DI, that measure similar concepts of physical function and disability.

Methods

Study design and participants

This study was a cross-sectional analysis of baseline data collected as part of a prospective observational study of outcome measures in IIM [12]. Patients were included who fulfilled either ACR/EULAR 2017 Classification or Bohan and Peter criteria for idiopathic inflammatory myopathies [15–17]. Patients were further characterized by IIM subtype: dermatomyositis, immune-mediated necrotizing myopathy (IMNM), polymyositis, anti-synthetase syndrome, or overlap myositis based on criteria followed at each individual site [18]. Patients were recruited from myositis centers in 5 countries (4 continents), including USA (Johns Hopkins Myositis Center), South Korea (Seoul National University Hospital), Netherlands (Amsterdam University Medical Center), Sweden (Karolinska University Hospital), and Australia (Perron Institute). Institutional Review Board approvals were obtained at participating centers and participants gave written informed consent in accordance with local regulations.

PROMIS measures

Three PROMIS measures for adults were included based on the proposed OMERACT core set for myositis: PROMIS Short Form v1.0—Pain Interference 6a [19], PROMIS Short Form v1.0—Fatigue 7a [20], and PROMIS Short Form v2.0—Physical Function 8b [21]. The PROMIS Pain Interference measure (short form 6a possible T-score range: 41.1–76.3) assesses the degree to which pain interferes with daily activities over the past week (higher scores indicate more interference due to pain). The PROMIS Fatigue measure (short form 7a possible T-score range: 29.4–83.2) assesses the impact of fatigue over the past week (higher scores indicate more fatigue). The PROMIS Physical Function measure (short form 8b possible T-score range: 20.3–60.1) assesses current ability to perform physical activities including housework/yardwork, climbing stairs, walking, and completing errands (in contrast to the other measures, higher scores indicate better physical function). All three measures are available in over 15 languages, including the dominant spoken languages at each clinical site: English, Korean, Dutch, and Swedish [22]. Participants completed the measures unassisted on paper or on a tablet computer. PROMIS measures are reported as T-scores with a population mean of 50 and standard deviation of 10, such that higher scores indicate more of the construct being measured. T-scores were calculated using the Health Measures Scoring Service (https://www.assessmentcenter.net/ac_scoringservice).

Myositis outcome measures

In addition to age, sex, diagnosis, and disease duration, five myositis core outcome measures were collected: patient and physician global disease activity on a visual analogue scale (0–10), manual muscle testing (MMT-8, range 0–80), serum creatine kinase (CK) activity within 6 weeks of the study visit, and Health Assessment Questionnaire Disability Index (HAQ-DI, range 0–3) as a measure of physical function that assesses activities including dressing/grooming, arising, eating, walking, hygiene, reaching, grip, and chores/errands (higher scores indicate worse function) [14]. HAQ-DI was collected at two sites (Netherlands

and Sweden). Given differences in measurement scales across countries, CK values were divided by the upper limit of normal for the assay to generate a value indicating the fold-change above the upper limit of normal.

Construct validity

Myositis Working Group voting members ($n = 14$), including two patient representatives, generated a priori hypotheses for the correlation between each of the 3 PROMIS measures (Pain Interference, Fatigue, Physical Function), patient age, and the 5 myositis core outcome measures (physician global, patient global, MMT-8, CK, and HAQ-DI). Consensus for the a priori hypotheses was pre-specified as greater than 70 % agreement among voting members on the expected strength of correlation, per OMERACT methodology [13]. An additional consensus meeting was held to refine a priori hypotheses when agreement was not reached after initial voting.

Statistical analysis

Patient characteristics, PROMIS measure T-scores, and disease outcome measures were summarized using descriptive statistics. Cronbach's alpha was used to assess internal consistency between items in each PROMIS measure with alpha >0.7 indicating acceptable internal consistency and alpha > 0.9 indicating high internal consistency [13, 23]. To demonstrate unidimensionality to support use of Cronbach's alpha for internal consistency, confirmatory factor analysis was used to evaluate each PROMIS measure, assuming single factor models [24]. Model fit statistics including comparative fit index (CFI)(threshold ≥0.9 adequate, ≥0.95 ideal), Tucker-Lewis Index (TLI) (threshold ≥0.9, ≥0.95 ideal), and the root mean squared error of approximation (RMSEA) (threshold <0.08) were calculated [25–27]. Floor and ceiling effects for each PROMIS measure were examined with a threshold of 15 % indicating presence of a floor or ceiling effect [28]. For correlation analyses, CK values were divided by the upper limit of normal for the assay to generate a value indicating the fold-change above the upper limit of normal, then log-transformation was applied due to a significantly right-skewed distribution. Pearson or Spearman's correlations between each PROMIS measure, the five myositis core outcome measures, and patient age were calculated. Pre-specified cut-offs for interpretation were weak $r < 0.4$, moderate $r = 0.4$–0.7, and strong $r > 0.7$ [29]. Pre-specified subgroup analysis according to IIM type (grouped as dermatomyositis and non-dermatomyositis IIM based on small group sizes for the non-dermatomyositis subtypes) was performed to compare correlation coefficients between groups. All analyses were performed using STATA 18 (StataCorp. 2023. *Stata Statistical Software: Release 18*. College Station, TX: StataCorp LLC.).

Results

Patient characteristics

This study included 135 patients with IIM (59 % female) from 5 countries, with median age 60 years (interquartile range (IQR): 50, 70) and median disease duration 1.2 years (IQR 0.05, 4.7). The most common IIM type was dermatomyositis with 65 (48 %) patients, followed by immune-mediated necrotizing myopathy with 29 (22 %) patients. Mean (SD) T-scores for PROMIS measures were Pain Interference 55.8 (11.5), Fatigue 57.0 (11.7), and Physical Function 39.3 (10.3). Additional characteristics and outcome measures are summarized in Table 1.

PROMIS measures

Distributions of PROMIS Pain Interference, Fatigue, and Physical Function T-scores are shown in Fig. 1. No floor or ceiling effects were observed for PROMIS Fatigue or Physical Function. Ceiling effect was observed for PROMIS Pain Interference T-scores with 37 (29 %) participants with the lowest score of 41.1 indicating low/no pain. These participants tended to have low disease activity: mean (SD) patient and physician global assessment of disease 2.8 (1.0, 5.0) and 2.0 (1.0, 6.0), respectively. Cronbach's alpha for PROMIS Pain Interference, Fatigue, and Physical Function were 0.98, 0.90, and 0.96, respectively, indicating high internal consistency reliability between items in each PROMIS measure. Confirmatory factor analysis supported that each of the PROMIS measures appeared to be unidimensional, loading onto only one factor. Model fit statistics were: PROMIS Pain Interference CFI 0.91, TLI 0.85, RMSEA (90 % confidence interval (CI)) 0.33 (0.29, 0.39); PROMIS Fatigue CFI 0.98, TLI 0.97, RMSEA (90 % CI) 0.09 (0.03, 0.13); PROMIS Physical Function CFI 0.92, TLI 0.88, RMSEA (90 % CI) 0.20 (0.17, 0.24).

Correlations between PROMIS measures and myositis outcome measures

A priori hypotheses and observed correlations are shown in Table 2. Consensus by voting members for a priori hypotheses was achieved for 15 of 18 hypotheses. After initial voting, additional member discussion led to consensus for a hypothesized moderate correlation between PROMIS Pain Interference and HAQ-DI. Consensus was not reached for the hypothesized correlation between PROMIS Fatigue and MMT (opinion divided between weak and moderate correlation), PROMIS Physical Function and creatine kinase (divided between weak and moderate correlation), or PROMIS Physical Function and HAQ-DI (divided between moderate and strong correlation). A priori hypotheses were confirmed for 5 of 6 (83 %) Pain Interference correlations, 4 of 5 (80 %) Fatigue correlations, and 3 of 4 (75 %) Physical Function correlations. Weak correlations were observed between age and each of the PROMIS measures.

For PROMIS Pain Interference, weak correlations were observed with physician global disease activity, MMT, CK, and HAQ-DI, while a moderate correlation was observed with patient global disease activity (Supplementary Fig. 1). For PROMIS Fatigue, weak correlations were observed with CK, while moderate correlations were observed with

Table 1
Characteristics of study patients.

	n (%) or mean (SD)
Number of participants	135
Country	
USA	40 (29.6 %)
South Korea	28 (20.7 %)
Netherlands	27 (20.0 %)
Sweden	24 (17.8 %)
Australia	16 (11.9 %)
Age (years)	59 (14)
Sex, female	79 (58.5 %)
Disease duration (years), median (IQR)	1.2 (0.05, 4.7)
Diagnosis	
Dermatomyositis	65 (48.1 %)
Immune-mediated necrotizing myopathy	29 (21.5 %)
Polymyositis	16 (11.9 %)
Anti-synthetase syndrome	16 (11.9 %)
Overlap myositis	9 (6.7 %)
Myositis outcome measures	
Patient global disease activity (0–10)	4.3 (2.8)
Physician global disease activity (0–10)	3.7 (2.6)
Manual muscle testing (0–80)	71.5 (11.2)
Creatine kinase relative to upper limit*, median (IQR)	0.9 (0.4, 3.8)
Health Assessment Questionnaire (HAQ-DI) (0–3) (n = 43)	1.2 (0.9)
PROMIS measure T-score	
PROMIS Pain Interference (n = 129)	55.8 (11.5)
PROMIS Fatigue (n = 135)	57.0 (11.7)
PROMIS Physical Function (n = 129)	39.3 (10.3)

IQR: interquartile range; SD: standard deviation; USA: United States of America.
* creatine kinase activity / upper limit of normal.

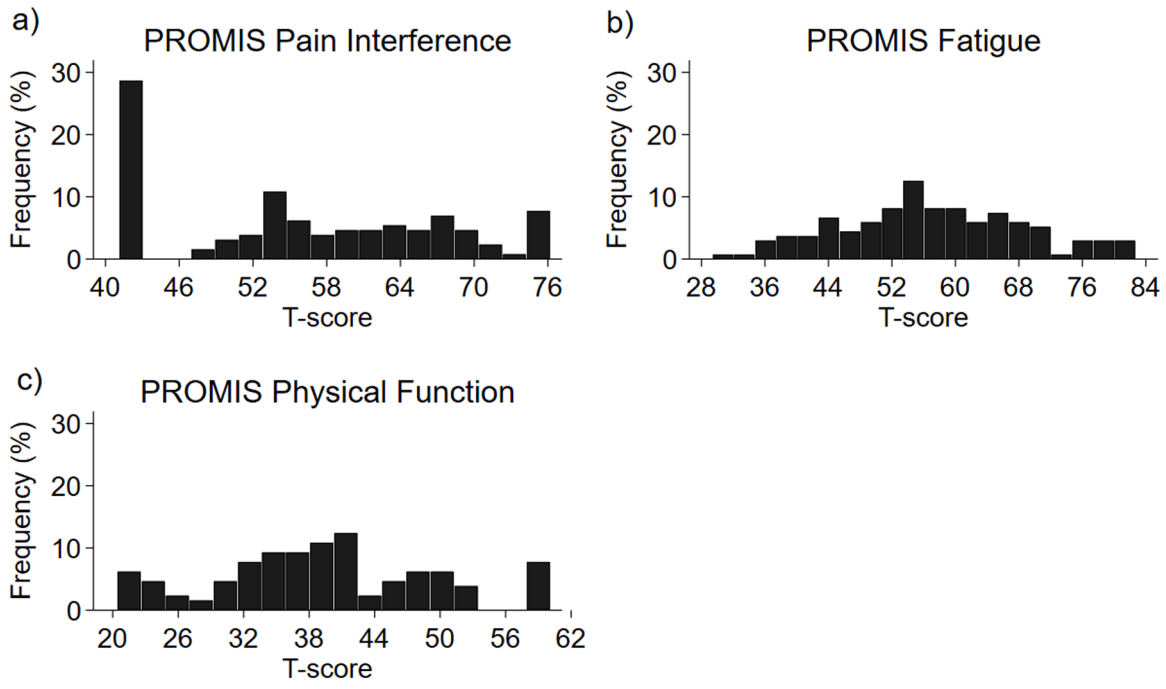

Fig. 1. Distribution of PROMIS T-scores.
Distribution of (a) PROMIS Short Form v1.0—Pain Interference 6a, (b) PROMIS Short Form v1.0—Fatigue 7a, and (c) PROMIS Short Form v2.0—Physical Function 8b T-scores. PROMIS Pain Interference demonstrates a ceiling effect with 29 % of patients recording the lowest possible score (low/no pain interference).

Table 2
A priori and observed correlations between PROMIS measures and myositis outcome measures.

	n	Consensus exercise, % agreement voting members ($n = 14$)	*A priori* hypothesis for expected correlation	Observed correlation		Hypothesis confirmed?
				Strength *	*r* (95 % CI)	
PROMIS Pain Interference						
Age	129	14/14 (100 %)	weak	weak	−0.14 (−0.3–0.03)	Yes
Patient global	104	11/14 (79 %)	moderate	moderate	0.46 (0.3–0.6)	Yes
Physician global	114	13/14 (93 %)	weak	weak	0.34 (0.2–0.5)	Yes
MMT	114	14/14 (100 %)	weak	weak	−0.21 (−0.4–0.0)	Yes
Creatine kinase	128	14/14 (100 %)	weak	weak	0.01 (−0.2–0.2)	Yes
HAQ-DI	40	10/14 (71 %)	moderate	weak	0.29 (0.0–0.6)†	No
PROMIS Fatigue						
Age	135	14/14 (100 %)	weak	weak	−0.16 (−0.3–0.0)	Yes
Patient global	108	13/14 (93 %)	moderate	moderate	0.54 (0.4–0.7)	Yes
Physician global	118	14/14 (100 %)	weak	moderate	0.51 (0.4–0.6)	No
MMT	119	8/14 (57 %)	weak	moderate	−0.41 (−0.5 to −0.2)	N/A
Creatine kinase	134	10/14 (71 %)	weak	weak	0.2 (0.0–0.4)	Yes
HAQ-DI	43	11/14 (79 %)	moderate	moderate	0.55 (0.3–0.7)†	Yes
PROMIS Physical Function						
Age	129	13/14 (93 %)	weak	weak	0.04 (−0.1–0.2)	Yes
Patient global	104	10/14 (71 %)	moderate	moderate	−0.64 (−0.7 to −0.5)	Yes
Physician global	114	11/14 (79 %)	weak	moderate	−0.56 (−0.7 to −0.4)	No
MMT	114	13/14 (93 %)	moderate	moderate	0.49 (0.3–0.6)	Yes
Creatine kinase	128	9/14 (64 %)	weak	weak	−0.18 (−0.3–0.0)	N/A
HAQ-DI	40	8/14 (57 %)	moderate	strong	−0.83 (−0.9 to −0.7)†	N/A

Creatine kinase = log of creatine kinase activity/upper limit of normal; CI: confidence interval; HAQ-DI: Health Assessment Questionnaire Disability Index; MMT: Manual muscle testing; N/A: not applicable, a priori consensus not reached so hypothesis not able to be confirmed.

* Cutoffs for interpretation were <0.4, 0.4–0.7, and >0.7 for weak, moderate, and strong correlations.

† Spearman correlation.

patient and physician global disease activity, MMT, and HAQ-DI (Supplementary Fig. 2). For PROMIS Physical Function, weak correlations were observed with CK; moderate correlations were observed with patient and physician global disease activity and MMT; strong correlation was observed with HAQ-DI (Supplementary Fig. 3). PROMIS measures showed moderate correlations with each other (Pain Interference and Physical Function $r = -0.6$, Pain Interference and Fatigue $r = 0.7$, Physical Function and Fatigue $r = -0.7$) (Supplementary Fig. 4).

Comparison of correlations for dermatomyositis vs. non-dermatomyositis IIM

Correlations between PROMIS measures and myositis outcome measures were examined for a subgroup of patients with dermatomyositis compared to non-dermatomyositis IIM types. Demographics, disease duration, core outcome measures, and PROMIS measures did not differ significantly between subgroups except for higher CK values in the non-dermatomyositis IIM subgroup, of which immune-mediated necrotizing myopathy was the most common diagnosis (Supplementary Table 1). The strength of correlations for individuals with dermatomyositis vs. non-dermatomyositis IIM were generally similar, although for PROMIS Physical Function and PROMIS Fatigue, correlations with each of the core measures except for HAQ-DI tended to be stronger in the non-dermatomyositis IIM group (Table 3). However, PROMIS Pain Interference and HAQ-DI were moderately correlated for the dermatomyositis group and weakly correlated for the non-dermatomyositis IIM group. This difference was attenuated but remained in a sensitivity analysis excluding patients with no or very low pain interference (i.e. ceiling effect) (dermatomyositis: $n = 10$, Spearman's $r = 0.81$; non-dermatomyositis IIM: $n = 18$, Spearman's $r = 0.39$). PROMIS Fatigue and physician global were weakly correlated for dermatomyositis and moderately correlated for non-dermatomyositis IIM. A similar pattern was observed for PROMIS Physical Function and both patient and physician global disease activity.

Table 3
Observed correlations between PROMIS measures and myositis outcome measures according to diagnosis.

	Observed *r* in full cohort	Dermatomyositis ($n = 65$)		Non-dermatomyositis [a] ($n = 70$)	
		n	Pearson's *r*	*n*	Pearson's *r*
PROMIS Pain Interference					
Age	−0.14	63	−0.16	66	−0.10
Patient global	0.46	52	0.49	52	0.44
Physician global	0.34	57	0.21	57	0.46
MMT	−0.21	57	−0.29	57	−0.14
Creatine kinase	0.01	62	0.02	66	0.06
HAQ-DI	0.29+	15	0.68+	25	0.05+
PROMIS Fatigue					
Age	−0.16	65	0.0	70	−0.3
Patient global	0.54	52	0.42	56	0.62
Physician global	0.51	57	0.31	61	0.64
MMT	−0.41	58	−0.38	61	−0.43
Creatine kinase	0.20	64	0.12	70	0.30
HAQ-DI	0.55+	15	0.65+	28	0.49+
PROMIS Physical Function					
Age	0.04	63	−0.02	66	0.10
Patient global	−0.64	52	−0.55	52	−0.73
Physician global	−0.56	57	−0.39	57	−0.69
MMT	0.49	57	0.39	57	0.57
Creatine kinase	−0.18	62	−0.10	66	−0.27
HAQ-DI	−0.83+	15	−0.91+	25	−0.76+

[a] Non-dermatomyositis includes immune-mediated necrotizing myopathy, polymyositis, anti-synthetase syndrome, and overlap myositis; +Spearman's rho.

Creatine kinase = log of creatine kinase level/upper limit of normal; HAQ-DI: Health Assessment Questionnaire Disability Index; MMT: Manual muscle testing; USA: United States of America.

Comparison of HAQ-DI and PROMIS physical function

As an exploratory analysis, we compared the floor/ceiling effect and construct validity of HAQ-DI and PROMIS Physical Function since both are outcome measures that intend to assess similar constructs. HAQ-DI was only available for a subset of patients (Netherlands and Sweden). To compare HAQ-DI and PROMIS Physical Function scores, only participants with both scores available were included for this analysis ($n = 40$). Similar to PROMIS Physical Function, the HAQ-DI did not demonstrate floor or ceiling effects (all <10 %). HAQ-DI and PROMIS Physical Function correlated strongly in the expected direction ($r = -0.83$). Correlations between HAQ-DI and the other core outcome measures were not significantly different compared to those observed for PROMIS Physical Function and the core outcome measures (Supplementary Table 2).

Discussion

This study demonstrates strong construct validity of PROMIS Short Form v1.0—Pain Interference 6a, PROMIS Short Form v1.0—Fatigue 7a, and PROMIS Short Form v2.0—Physical Function 8b measures in comparison to IMACS core outcome measures in an international cohort of adult patients with IIM, excluding inclusion body myositis. Compared to a prior study, we re-demonstrate high internal consistency of PROMIS items for each measure using a separate cohort [11]. We also demonstrate a very similar range and distribution of PROMIS T-scores for each measure compared to prior work in a separate cohort, indicating consistency of these results in adult IIM [11]. Our subgroup analysis according to IIM subtype, grouped as dermatomyositis and non-dermatomyositis IIM based on small group sizes for the non-dermatomyositis subtypes, shows generally consistent results for dermatomyositis and non-dermatomyositis IIMs compared to the full cohort.

The construct validity for PROMIS Pain Interference was strong with 5 of 6 (83 %) a priori hypotheses confirmed. The a priori hypothesis of a moderate correlation for PROMIS Pain Interference and HAQ-DI was not met, with results demonstrating a weak correlation ($r = 0.29$). The working group initially did not achieve consensus on this comparison, with opinions divided between weak and moderate, but agreed upon moderate after discussion. However, the prevailing group opinion was that of a 'weak-moderate' correlation, acknowledging that significant pain would limit physical function. Similar to prior work [11], PROMIS Pain Interference was demonstrated to have a ceiling effect in which approximately one-third of patients reported little or no pain. This finding was observed in every IIM subtype in this cohort and associated with lower disease activity. While it might be possible to reduce this observed effect with Computer Adaptive Testing (CAT) or a longer test form such as the 8-item short form, this may not be necessary. It is likely that a proportion of patients will continue to have the best possible score, not indicating a true ceiling effect of the instrument, but, rather, that some patients do not have pain.

The construct validity of PROMIS Fatigue was strong with 4 of 5 (80 %) a priori hypotheses confirmed. The hypothesis was not confirmed for PROMIS Fatigue and physician global, with the group expecting a weak correlation while a moderate correlation was observed. Working group members did not reach consensus regarding the proposed association between PROMIS Fatigue and MMT, with votes divided between weak and moderate correlations. Our results showed a correlation of −0.41, close to the threshold between weak and moderate. Discussion among

working group members highlighted that fatigue can be independent of weakness, though patient representatives also felt that fatigue itself may lead to worse performance on the MMT. The scatterplot of PROMIS Fatigue vs. MMT supports these experiences, demonstrating high fatigue scores across the full range of MMT scores (see Supplementary Fig. 2).

The construct validity of PROMIS Physical Function was good with 3 of 4 (75 %) a priori hypotheses confirmed. As with PROMIS Fatigue, the hypothesis for PROMIS Physical Function and physician global was not met due to a hypothesized weak correlation but observed moderate correlation. Voting members did not reach consensus for the proposed association between PROMIS Physical Function and serum creatine kinase value, though the majority hypothesized a weak correlation, which is what was observed. In discussion, it was felt that this association may depend on myositis type. For example, immune-mediated necrotizing myopathy is often associated with high CK and may have a stronger correlation between CK and PROMIS Physical Function compared to dermatomyositis with more modest CK elevations, but additional studies would be needed to determine this possibility since median CK values were not that abnormal in this cohort. The group also did not reach consensus on the association between PROMIS Physical Function and HAQ-DI, with members divided between moderate and strong correlations based on prior work demonstrating moderate-strong correlation between a measure of physical function (Myositis Activities Profile) and HAQ-DI [30]. Our results demonstrated strong correlation between PROMIS Physical Function and HAQ-DI.

We performed a subgroup analysis to compare observed correlations between patients with dermatomyositis and patients with non-dermatomyositis IIMs. Although our study was not powered to detect differences between correlations across IIM subgroups, comparison of the *rho* values demonstrates that many correlations were similar to the primary analysis, suggesting that these PROMIS measures have construct validity across IIM subtypes. A few notable differences include a moderate correlation between PROMIS Pain Interference and HAQ-DI in dermatomyositis but very weak correlation between these two measures for other IIMs. This analysis is limited by the small sample size but suggests that the relationship between pain and physical function may differ by IIM subtype. The associations between physician global and PROMIS Fatigue and PROMIS Physical Function, respectively, were weak in dermatomyositis but moderate in non-dermatomyositis IIMs, suggesting that physician incorporation of physical function may weigh differently when rating global disease activity across IIM subtypes, though further investigation would be needed to evaluate this. Interestingly, these two correlations (PROMIS Fatigue with physician global and PROMIS Physical Function with physician global) were two where the working group's a priori hypotheses were not met, which may reflect the influence of these differences between IIM subtypes.

As an exploratory analysis we compared correlations between HAQ-DI or PROMIS Physical Function and the core outcome measures and found that correlations with the core outcome measures were similar between HAQ-DI and PROMIS Physical Function, supporting that these measures assess similar constructs. Since the HAQ-DI contains 20 questions compared to 8 in the PROMIS Physical Function short form, the latter may reduce burden to the patient. However, despite overlap, the question content varies between the measures with greater emphasis on activities of daily living and distal motor functions (i.e. grip, eating, grooming) in the HAQ-DI, for which content validity has not been directly assessed in IIM, compared to global function (i.e. vacuuming, climbing stairs, running errands) in the PROMIS Physical Function measure. Thus, these instruments may provide complementary information, but further studies with larger sample sizes would be beneficial.

We observed Cronbach's alpha above the threshold for good internal consistency reliability for all measures, and above 0.95 for two measures (PROMIS Pain Interference and Physical Function), suggesting possible redundancy of item information. However, no adjustments are suggested to these well-established measures, which as short forms present low burden to patients. In performing confirmatory factor analysis, each of the PROMIS measures appeared to be unidimensional, supporting our results for internal consistency using Cronbach's alpha. The CFIs were greater than 0.9, indicating acceptable incremental fit but the RMSEs for Pain Interference and Physical Function were >0.1, indicating poor absolute fit. While this analysis is limited by small sample size and is not intended to fully evaluate the structural validity of these measures, these findings may indicate that these specific PROMIS measures do not have ideal psychometric properties for the myositis study population for which these statistics were calculated. However, we chose prespecified fixed/short forms for ease of administration, which cannot be deconstructed. Future iterations of our work may explore computer adaptive tests from the PROMIS Pain Interference and Physical Function item banks, which may improve model fit.

This study has several strengths including the use of an international cohort of adult IIM with representation from multiple continents and non-English speaking countries, supporting the generalizability of these findings. This study provides data regarding the construct validity of the selected OMERACT core set PROs in comparison to core outcome measures for disease activity in IIM and across IIM subtypes. Our results demonstrate consistency with prior work in separate cohorts, supporting the utility of these PROMIS measures in IIM. Some limitations of this work include small sample sizes and limited data availability regarding disease manifestations (extra-muscular features, auto-antibodies) and treatments. HAQ-DI was recorded in two of the five participating centers, limiting the comparison of HAQ-DI with other measures. Lastly, the working group was not able to reach consensus on 3 of the proposed correlations, which reduces the number of correlations used to assess construct validity. However, the working group discussions including patient input helped to clarify the areas of uncertainty regarding the relationship between these PROMIS measures and core outcome measures.

Conclusion

In conclusion, this study demonstrates strong construct validity of PROMIS Short Form v1.0—Pain Interference 6a, PROMIS Short Form v1.0—Fatigue 7a, and PROMIS Short Form v2.0—Physical Function 8b in comparison to myositis core outcome measures in an international adult IIM cohort and demonstrates similar findings across IIM subtypes. These results add to the growing body of work supporting the validity of these PROs as the selected instruments to assess core domains of life impact in adult IIM in both clinical practice and clinical trials.

Grant or other funding support

This study was supported by NIAMS Rheumatology Research Training Grant T32-AR076951-04 (Dr. Romich) and NIAMS K23AR075898 (Dr. Mecoli).

CRediT authorship contribution statement

Ellen Romich: Conceptualization, Formal analysis, Writing – original draft. **Didem Saygin:** Conceptualization, Formal analysis, Investigation, Writing – original draft. **Dana DiRenzo:** Conceptualization, Formal analysis, Investigation, Writing – original draft. **Christopher A. Mecoli:** Conceptualization, Formal analysis, Investigation, Writing – original draft. **Ingrid de Groot:** Investigation, Writing – review & editing. **Karin Lodin:** Investigation, Writing – review & editing. **Malin Regardt:** Investigation, Writing – review & editing. **Catherine Sarver:** Investigation, Writing – review & editing. **Ju Yeon Kim:** Investigation, Writing – review & editing. **Jin Kyun Park:** Investigation, Writing – review & editing. **Kelly Beer:** Investigation, Writing – review & editing. **Merrilee Needham:** Investigation, Writing – review & editing. **Helene Alexanderson:** Investigation, Writing – review & editing. **Lisa Christopher-Stine:** Investigation, Writing – review & editing. **Marianne de Visser:** Conceptualization, Data curation. **Joost Raaphorst:**

Investigation, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Merrilee Needham has received honorarium for educational talks and advisory boards from CSL, Sanofi-Aventis, and Teva.

Ellen Romich received support from NIH Rheumatology Research Training Grant T32-AR076951 which was paid to the institution.

Christopher A. Mecoli received support from K23AR075898 Grant.

Acknowledgments

We wish to acknowledge the guidance from Beverley Shea, Ph.D. and Dorcas Beaton, Ph.D. with the OMERACT technical advisory group. The Dutch authors wish to acknowledge Caroline Holtslag's assistance in data acquisition and entry.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.semarthrit.2024.152534.

References

[1] Lundberg IE, Fujimoto M, Vencovsky J, et al. Idiopathic inflammatory myopathies. Nat Rev Dis Prim 2021;7(1):87. https://doi.org/10.1038/s41572-021-00325-7.
[2] Bronner IM, van der Meulen MFG, de Visser M, et al. Long-term outcome in polymyositis and dermatomyositis. Ann Rheum Dis 2006;65(11):1456–61. https://doi.org/10.1136/ard.2005.045690.
[3] Armadans-Tremolosa I, Selva-O'Callaghan A, Visauta-Vinacua B, Guilera G, Pinal-Fernández I, Vilardell-Tarrés M. Health-related quality of life and well-being in adults with idiopathic inflammatory myopathy. Clin Rheumatol 2014;33(8):1119–25. https://doi.org/10.1007/s10067-014-2685-0.
[4] Maxwell LJ, Beaton DE, Boers M, et al. The evolution of instrument selection for inclusion in core outcome sets at OMERACT: filter 2.2. Semin Arthritis Rheum 2021;51(6):1320–30. https://doi.org/10.1016/j.semarthrit.2021.08.011.
[5] OMERACT working group: myositis. Accessed April 4, 2024. https://omeract.org/working-groups/myositis/.
[6] Park JK, Mecoli CA, Alexanderson H, et al. Advancing the development of patient-reported outcomes for adult myositis at OMERACT 2016: an international delphi study. J Rheumatol 2017;44(11):1683–7. https://doi.org/10.3899/jrheum.161252.
[7] 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(1):106–11. https://doi.org/10.3899/jrheum.180353.
[8] Regardt M, Mecoli CA, Park JK, et al. OMERACT 2018 modified patient-reported outcome domain core set in the life impact area for adult idiopathic inflammatory myopathies. J Rheumatol 2019;46(10):1351–4. https://doi.org/10.3899/JRHEUM.181065.
[9] Maxwell LJ, Beaton DE, Shea BJ, et al. Core domain set selection according to OMERACT Filter 2.1: the OMERACT methodology. J Rheumatol 2019;46(8):1014–20. https://doi.org/10.3899/jrheum.181097.
[10] Esfandiary T, Park JK, Alexanderson H, et al. Assessing the content validity of patient-reported outcome measures in adult myositis: a report from the OMERACT myositis working group. Semin Arthritis Rheum 2020;50(5):943–8. https://doi.org/10.1016/J.SEMARTHRIT.2020.06.006.
[11] DiRenzo D, Saygin D, de Groot I, et al. Reliability and validity of PROMIS physical function, pain interference, and fatigue as patient reported outcome measures in adult idiopathic inflammatory myopathies: international study from the OMERACT myositis working group. Semin Arthritis Rheum 2023;58. https://doi.org/10.1016/J.SEMARTHRIT.2022.152111.
[12] Saygin D, DiRenzo D, Raaphorst J, et al. Responsiveness and meaningful thresholds of PROMIS pain interference, fatigue, and physical function forms in adults with idiopathic inflammatory myopathies: report from the OMERACT myositis working group. Semin Arthritis Rheum 2024;64:152339. https://doi.org/10.1016/j.semarthrit.2023.152339.
[13] Beaton D., Maxwell L., Grosskleg S., et al., editors *The OMERACT HANDBOOK*. Version 2. OMERACT Accessed March 27, 2024. https://omeract.org/handbook/.
[14] Rider LG, Giannini EH, Harris-Love M, et al. Defining clinical improvement in adult and juvenile myositis. J Rheumatol 2003;30(3):603–17.
[15] Lundberg IE, Tjärnlund A, Bottai M, et al. 2017 European league against rheumatism/american college of rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. Ann Rheum Dis 2017;76(12):1955–64. https://doi.org/10.1136/annrheumdis-2017-211468.
[16] Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). N Engl J Med 1975;292(7):344–7. https://doi.org/10.1056/NEJM197502132920706.
[17] Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). N Engl J Med 1975;292(8):403–7. https://doi.org/10.1056/NEJM197502202920807.
[18] Leclair V, Lundberg IE. New myositis classification criteria-what we have learned since Bohan and Peter. Curr Rheumatol Rep 2018;20(4):18. https://doi.org/10.1007/s11926-018-0726-4.
[19] Amtmann D, Cook KF, Jensen MP, et al. Development of a PROMIS item bank to measure pain interference. Pain 2010;150(1):173–82. https://doi.org/10.1016/j.pain.2010.04.025.
[20] Lai J-S, Cella D, Choi S, et al. How item banks and their application can influence measurement practice in rehabilitation medicine: a PROMIS fatigue item bank example. Arch Phys Med Rehabil 2011;92(10 Suppl):S20–7. https://doi.org/10.1016/j.apmr.2010.08.033.
[21] Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware JEJ. The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. J Clin Epidemiol 2014;67(5):516–26. https://doi.org/10.1016/j.jclinepi.2013.10.024.
[22] PROMIS Health Measures. Accessed April 4, 2024. https://www.healthmeasures.net/explore-measurement-systems/promis.
[23] Cronbach L.J. My current thoughts on coefficient alpha and successor procedures. Shavelson R.J., editor *Educ psychol meas* 2004;64(3):391–418. doi:10.1177/0013164404266386.
[24] Cortina JM. What is coefficient alpha? An examination of theory and applications. J Appl Psychol 1993;78(1):98–104. https://doi.org/10.1037/0021-9010.78.1.98.
[25] Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990;107(2):238–46. https://doi.org/10.1037/0033-2909.107.2.238.
[26] Hooper D, Coughlan J, Mullen M. Structural equation modeling: guidelines for determining model fit. Electron J Bus Res Methods 2007;6.
[27] Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res Int J Qual Life Asp Treat Care Rehabil 2018;27(5):1147–57. https://doi.org/10.1007/s11136-018-1798-3.
[28] Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60(1):34–42. https://doi.org/10.1016/J.JCLINEPI.2006.03.012.
[29] Dancey CP, Reidy J. Statistics without maths for psychology. Pearson Education; 2007.
[30] 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(11):2386–92.