



Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

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

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

Dana DiRenzo^{a,1}, Didem Saygin^{n,1}, Ingrid de Groot^g, Clifton O. Bingham III^b, Ingrid E. Lundberg^h, Merrilee Needham^f, Jin Kyun Park^c, Malin Regardt^e, Catherine Sarver^m, Yeong Wook Song^j, Lara Maxwell^k, Dorcas Beaton^l, Marianne de Visserⁱ, Lisa Christopher-Stine^b, Christopher A. Mecoli^{b,2,*}, Helene Alexanderson^{d,2}

^a Division of Rheumatology, Department of Medicine, University of Pennsylvania, USA

^b Division of Rheumatology, Department of Medicine, Johns Hopkins University, USA

^c Division of Rheumatology, Department of Internal Medicine, Seoul National University Hospital and College of Medicine, Seoul, Korea

^d Physiotherapist, Medical Unit Occupational Therapy and Physical Therapy, Karolinska University Hospital and Division of Rheumatology, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden

^e Occupational Therapist, 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

^f Dept Neurology, Fiona Stanley Hospital, IHD Murdoch University and University of Notre Dame, Perth, Australia

^g Patient Research Partner, the Netherlands

^h Division of Rheumatology, Department of Medicine, Karolinska Institutet, Karolinska University, Solna, Stockholm, Sweden

ⁱ Department of Neurology, Amsterdam University Medical Centers, location Academic Medical Center, University of Amsterdam, the Netherlands

^j Division of Rheumatology, Department of Internal Medicine, Medical Research Center, College of Medicine, Department of Molecular Medicine and Biopharmaceutical Sciences, Seoul National University

^k Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

^l Institute for Work & Health and Institute for Health Policy Management and Evaluation, University of Toronto, Toronto, Ontario, Canada

^m Patient Research Partner, USA

ⁿ Division of Rheumatology, Department of Medicine, University of Chicago, USA

ABSTRACT

Objective: Pain interference, fatigue, and impaired physical function are common features of idiopathic inflammatory myopathies (IIM). The objective of this study was to evaluate the construct validity and test-retest reliability of the Patient Reported Outcome Information System (PROMIS) Pain Interference 6av1.0, Fatigue 7av1.0, and Physical Function 8bv2.0 instruments.

Methods: Patient-Reported Outcome Measures (PROMs) were deployed to adult IIM patients from OMERACT Myositis Working Group (MWG) international clinic sites via two online surveys (2019, 2021). Internal consistency of each PROM was analyzed by Cronbach's α . Construct validity was determined by *a priori* hypotheses generated by the MWG with >75% agreement for each hypothesis and calculated with Pearson correlations. Test-retest reliability was assessed using intraclass correlation coefficient with PROMIS instruments administered at time zero and 7 days.

Results: Surveys were sent to 368 participants in total; participants who completed each questionnaire varied (n=65 to 263). For construct validity, 10 out of 13 *a priori* hypotheses were met supporting construct validity of PROMIS instruments (Pain Interference 3/4, fatigue 4/4, and Physical Function 3/5). Test-retest reliability was strong for all PROMIS instruments. All PROMIS instruments demonstrated excellent internal consistency. None of the measures demonstrated any ceiling or floor effects except for a ceiling effect in the Pain Interference instrument.

* Corresponding author at: Johns Hopkins University School of Medicine, Division of Rheumatology, 5200 Eastern Ave, MFL Bldg, Center Tower, Suite 4100, Baltimore, MD 21224, USA.

E-mail address: cmecoli1@jhmi.edu (C.A. Mecoli).

¹ Co-First Authors.

² Co-Senior Authors.

<https://doi.org/10.1016/j.semarthrit.2022.152111>

Available online 3 November 2022

0049-0172/© 2022 Elsevier Inc. All rights reserved.

Conclusions: This study presents test-retest reliability and construct validity evidence supporting PROMIS Pain Interference (6a v1.0), Fatigue (7a v1.0), and Physical Function (8b v2.0) using a large international cohort of patients with IIM. Internal consistency of these instruments was excellent. A ceiling effect was noted in the Pain Interference instrument.

Introduction

Idiopathic inflammatory myopathies (IIM, colloquially known as ‘myositis’) are a rare group of heterogeneous systemic autoimmune diseases including polymyositis, dermatomyositis, immune mediated necrotizing myopathy, anti-synthetase syndrome, and overlap myositis. Although muscle weakness is the cardinal feature of IIM, extra-muscular manifestations including constitutional symptoms, interstitial lung disease, rash, arthritis, cardiac and cutaneous manifestations, can also be seen [1]. Despite favorable effects of available therapies, a majority of the patients develop sustained limitations in daily activities and quality of life [2,3].

Currently available outcome measures in IIM primarily focus on healthcare provider evaluation and assessment of disease activity. While these outcome measures are a critically important component of patient assessment, they do not adequately capture the symptoms frequently experienced by patients including pain, fatigue, and reduced physical function [4]. The importance of incorporating the patient perspective into clinical decision-making has been increasingly recognized over the past several decades and has also been encouraged by federal agencies and medical societies [5]. Despite the acknowledged importance of the patient perspective, there are currently no validated measures of fatigue or pain in myositis. Studies frequently use visual analog scales for pain and fatigue, which give a rough quantification of the level of pain and fatigue experienced by the patient on a 10-point scale and are easy-to-use in a busy clinical setting; however, these do not provide further information on the effect of pain and fatigue on daily activities or quality of life [6]. The Medical Outcomes Study Questionnaire Short Form 36 (SF36) pain and fatigue subscales have also been used in previous studies; however, psychometric properties of these subscales remain to be studied in patients with adult IIM and were neither developed nor validated for independent use [4]. Patient-reported outcome measures (PROMs) commonly used in IIM (e.g., Health Assessment Questionnaire, Patient Global Assessment, SF-36), while useful, have several limitations including a lack of data on content validity [7]. There is thus a need for patient-derived, patient-reported, widely available, reliable, and valid outcome measures for use in myositis clinics and clinical trials.

Recognizing this important need, Outcome Measures in Rheumatology (OMERACT) brings together multiple stakeholders to facilitate the development and/or validation of PROMs under an established framework [8]. The OMERACT Myositis Working Group, established in 2011, is a multidisciplinary international group with members from Australia, Canada, South Korea, Sweden, the Netherlands, and the United States, and consists of healthcare providers, methodologists, and patient research partners. Over the past decade, several focus groups were conducted in three countries (USA, Sweden, and South Korea) to investigate patients’ experience of living with myositis to identify the symptom domains that matter most [9,10]. Transcripts of the focus groups were analyzed qualitatively with identification of 26 candidate domains [11]. This was followed by three rounds of international modified Delphi exercises with participation of patients, health care providers, and caregivers from three continents to determine the core domains of interests [11–13]. Fatigue, pain interference, and physical activity (later operationalized as physical function based on focus groups) were identified as three core domains of life impact that should be prioritized to assess in clinical trials [12,13]. After determination of the domains of interest, candidate instruments were identified based on literature review and Myositis Working Group discussions [7,14]. Final

candidate instruments were selected based on cognitive debriefing with patient groups regarding comprehensibility, feasibility, clarity, and content validity of each candidate instrument: PROMIS physical function short form 8b, PROMIS pain interference short form 6a (v1), and PROMIS fatigue short form 7a [14].

In this study, our objectives were to: [1] provide formal definitions of these three core domains, [2] assess construct validity, and [3] test-retest reliability of the PROMIS fatigue, physical function and pain measures in a large international cohort of adult IIM patients.

Methods

Patients

Patients with probable or definite dermatomyositis/polymyositis (according to Bohan-Peter and/or 2017 ACR/EULAR Classification Criteria) and immune-mediated necrotizing myopathy according to the 2003 European Neuromuscular Centre criteria were included [15–17]. Patients with inclusion body myositis were excluded from the study. Clinical care sites included The Johns Hopkins Division of Rheumatology, Baltimore, MD, United States, the Karolinska University Hospital, Stockholm, Sweden, the Amsterdam University Medical Centers, Amsterdam, the Netherlands, the Seoul National University Hospital in Seoul, South Korea, and the Fiona Stanley Hospital in Murdoch, Australia. This study was approved by the respective ethic review boards as obligated [Johns Hopkins Hospital (IRB number: IRB00098790); Karolinska Institutet (2017/1697-31); Amsterdam University (W20_320 # 20.356); Seoul National University Hospital (1312-009-537), and Australia (Murdoch University HREC Approval 2015/111)].

Candidate instruments

Formalized definitions of each domain (fatigue, pain interference, physical function) were developed per methodological guidance from the OMERACT Instrument Selection Filter 2.1, the OMERACT Handbook as previously described [18,19] [Table 1]. PROMIS fatigue 7a v1.0 (7 items), PROMIS pain interference 6a v1.0 (6 items), and PROMIS physical function 8b v2.0 (8 items) were determined to be representative instruments for the domains of interest based on prior exercises conducted by the OMERACT Myositis Working Group [14]. Of note, while PROMIS pain interference 8a was originally reviewed by our patient focus groups because of its excellent reliability and validity in rheumatoid arthritis, this instrument was not available across multiple

Table 1

Definitions and selected instruments for the OMERACT myositis working group core domains of interest.

OMERACT DOMAIN	Definition	Selected Instrument
Physical Function	The ability to perform basic and desired activities of daily living that is affected by decreased use of muscles or other organ systems affected by IIM.	PROMIS Physical Function 8b v2.0
Pain Interference	Aching, soreness, or tenderness attributable to IIM relating to the joints, muscles, and/or skin causing interference with routine daily function.	PROMIS Pain Interference 6a v1.0
Fatigue	A feeling of [extreme] tiredness or exhaustion attributable to IIM, interfering with usual and meaningful daily activities.	PROMIS Fatigue 7a v1.0

languages (see below) [20]. Thus, PROMIS pain interference 6a v1.0 (6 items) was selected by OMERACT technical advisory group and approved by the Myositis Working Group. The PROMIS 6a is identical to the PROMIS 8a instrument except for two additional questions contained in the 8a: “How much did pain interfere with your enjoyment in life?” (PAININ3) and “How much did pain interfere with your family life?” (PAININ13).

For each PROMIS instrument, each question has 5 responses, ranging in value from 1-5. Raw question scores are totaled, converted to T-scores (calibrated, and normalized to a population mean score of 50 with a standard deviation of 10) [21]. Higher scores indicate more of the concept being measured. T-scores range from 41.1-76.3 for PROMIS pain interference 6a v1.0, 29.4-83.2 for PROMIS fatigue 7a v1.0, and 20.3-60.1 for PROMIS physical function 8b v2.0. All three instruments utilize a 7-day recall period.

Language translation

The three instruments of interest are available in over 15 languages, including the dominant, spoken languages at each clinical site: English, Korean, Dutch, and Swedish [22].

Construct validity

For construct validity, a total of 14 *a priori* hypotheses were generated based on consensus (*i.e.*, >75% agreement by the Myositis Working Group members required for each hypothesis). Failure to achieve >75% agreement resulted in the hypothesis being dropped.

PROM data were extracted and analyzed from two different sources: the OMERACT Myositis Working Group 2019 Survey of PROM Content Validity and Feasibility, and the OMERACT Myositis Working Group 2021 Survey of PROM Construct Validity and Reliability. Of note, the 2021 Survey was administered concurrently with the test-retest exercise, thus participants are identical. The two surveys included questions on demographics as well as the following PROMs for assessment of construct validity of the candidate instruments (PROMIS physical function, fatigue, and pain interference): Patient Global Assessment of Disease Activity (0-10, 0 equals no activity), the Pain Disability Index (PDI, range 0 to 10, 0 equals no disability), pain intensity numeric rating scale, International Physical Activity Questionnaire (IPAQ) (Total METS/Mins/Week subset), Myositis Activities Profile (MAP) (range 1 to 7, 1 equals no limitation), and PROMIS 4a v1.0 anxiety (T-Score, 40.3-81.6), depression (T-Score, 41.0-79.4), sleep disturbance (T-Score, 32.0-73.3), and social participation (T-Score, 27.9-63.8) questionnaires [23–26]. Pearson or Spearman correlation was then calculated to assess correlation between three candidate instruments (PROMIS physical function, fatigue, and pain interference) and other measures as appropriate depending on the normality of the data distribution. Evaluation of weak ($r < 0.25$), moderate ($r \geq 0.25$ to < 0.75), and strong ($r \geq 0.75$) correlations, was determined by group consensus and broadly guided by previous literature [27]. Student's *t*-test was used to compare the results of the PROMIS instruments between male vs females. All analyses were performed using Stata version 15.1 (College Station, TX).

Test-retest reliability

The test-retest reliability assessment was deployed via Redcap, an online data management platform, to all adult patients with valid email addresses at the North American clinical care site (Johns Hopkins University, Baltimore, MD). PROMIS physical function, pain interference and fatigue instruments were distributed at baseline (time 0) and again 7 days later. Each PROMIS item has 7-day recall period. Therefore, based on the working group consensus, a test-retest period of 7 days was predicted to be clinically stable period for IIM. The pain numeric rating scale (NRS), (range 0 to 10 with 10 representing the most severe pain), was used as an anchor and a paired *t*-test was performed to ensure

stability. Test-retest reliability was calculated with two-way random-effects intraclass correlation coefficients (ICC) and Bland-Altman plots. ICC was interpreted as poor reliability if $ICC < 0.5$, moderate reliability if between 0.5-0.75, and good reliability if > 0.75 [28]. Correlation matrices were constructed as well as histogram plots of baseline T-Scores for each instrument. Internal consistency was assessed with Cronbach alpha, which was interpreted as good if 0.8-0.9, and excellent with possible redundancy if ≥ 0.9 [29].

Results

Domain definitions

Definitions for the three core domains were formalized based on Myositis Working Group consensus and on prior work [14] (Table 1).

Study cohort

The number of participants varied based on which survey was completed (2019 or 2021, Tables 2 and 3). For the 2019 survey, 368 patients initiated the survey. Of the 368, only a subgroup completed the entire survey, including the demographic portion as well as all individual instruments. The mean (SD) age was 60 (11) and 73% were female. Eighty percent of participants were from the United States ($n=129$), followed by Australia ($n=8$, 5%), the United Kingdom ($n=7$, 4%), Canada ($n=5$, 3%), the Netherlands ($n=2$, 1%), and Sweden ($n=2$, 1%). The mean (SD) patient global assessment of disease activity (PGA, 0-10, 0=no disease activity) was 5.1 (2.6), PROMIS pain interference T-Score was 58.7 (11.3), fatigue T-Score was 61.9 (9.2), physical function T-Score was 37.1 (6.9), IPAQ Total METS/Mins/Week was 4767.6 (3517), MAP score was 3.0 (1.8), and PDI mean sum was 29.8 (20.8) (Fig. 2). The 2021 survey had a similar demographic composition and was completed by the same participant pool as the test-retest exercise ($n=263$). Participants were 76% female with a mean (SD) age of 60 (12). The mean (SD) baseline PROMIS pain interference (53.6 (9.9)) and fatigue T-Scores (54.7 (10.2)) fell within $\frac{1}{2}$ standard deviation of US population normative values. However, the PROMIS physical function mean T-Score (SD) was nearly one standard deviation worse than the population mean at 41.7 (9.2).

Construct validity

For the construct validity exercise, a total of 14 hypotheses were developed *a priori* by the Myositis Working Group. Consensus was achieved (>75% agreement) for 13 of the 14 hypotheses. In Myositis Working Group discussions, there was a <75% agreement regarding the relationship between PROMIS pain interference and sex, thus this item was dropped from the exercise. For the remaining 13 hypotheses, 3 out of the 4 hypotheses were met for PROMIS pain interference, 4 out of 4 for fatigue, and 3 of 5 for physical function tools (Table 3).

The PROMIS physical function instrument correlated strongly with PROMIS social participation instrument, moderately with the PROMIS anxiety, depression, and sleep disturbance instruments, and the pain NRS, as well as patient global disease activity. The PROMIS physical function instrument correlated poorly with age (Table 3, Supplementary Table 1). The PROMIS physical function scores were similar between males and females ($p=0.64$). The PROMIS physical function scores had high positive correlation with the MAP subscales of movement ($r=0.82$), housework ($r=0.84$), activities of moving around ($r=0.76$), and personal care ($r=0.78$) (Supplementary Table 2). Interestingly, there was weak negative correlation between the PROMIS physical function and IPAQ total METS/min/week ($r=-0.26$). Generally, the IPAQ METS/min/week did not correlate strongly with any of the MAP subscales or PROMIS physical function (r for all comparisons < -0.3).

The PROMIS pain interference instrument correlated strongly with the pain NRS, moderately with the PROMIS fatigue, physical function,

Table 2
Mean PROMIS T-scores of the baseline and follow-up visits and ICC for test-retest reliability.

	Baseline (Test)		Follow-up (Re-Test)		Mean Difference (95% CI)	ICC(95% CI)
	n	Mean T-Score (SD)	n	Mean T-Score (SD)		
PROMIS Instruments						
Physical Function	272	41.7 (9.2)	159	41.4 (9.0)	-0.3 (-2.1, 1.5)	0.97 (0.96, 0.98)
Pain Interference	279	53.6 (9.9)	164	53.3 (9.7)	-0.3 (-2.2, 1.6)	0.93 (0.91, 0.95)
Fatigue	275	54.7 (10.2)	162	53.2 (10.3)	-1.5 (-3.5, 0.4)	0.94 (0.91, 0.95)
Anchor						
Pain NRS	271	3.0 (2.6)	159	3.0 (2.6)	-0.1 (-0.6, 0.4)	0.95 (0.93, 0.96)

SD: Standard Deviation, CI: Confidence interval, ICC: Intraclass Correlation Coefficient, NRS: Pain Numeric Rating Scale.

Table 3
Correlations between the PROMIS physical function, pain interference, and fatigue, forms, and other variables and instruments according to the *a priori* hypotheses generated by the Working Group.

Variables	n	<i>A priori</i> hypothesis- expected correlation	% Agreement of Members	Pearson's r	Hypothesis confirmed
PROMIS Physical Function					
PROMIS Depression (4a v1)	263	Moderate	93%	-0.49	Yes
Sex	78	Low	100%	T test*, p=0.6	Yes
Age	78	Low	100%	-0.11	Yes
International Physical Activity Questionnaire	100	Moderate	86%	-0.13	No
Myositis Activities Profile	100	Moderate	86%	0.77	No
PROMIS Pain Interference					
PROMIS Sleep Disturbance (4a v1)	263	Moderate	93%	0.35	Yes
Pain Intensity Numeric Scale	263	Moderate	100%	0.87	Yes
Sex	N/A	Low	57%	Dropped	N/A
Age	65	Low	93%	-0.18	Yes
Pain Disability Index	100	High	100%	0.63	No
PROMIS Fatigue					
PROMIS Anxiety (4a v1.0)	263	Moderate	93%	0.58	Yes
PROMIS Sleep Disturbance (4a v1)	263	Moderate	100%	0.54	Yes
Sex	68	Low	100%	T test*, p=0.6	Yes
Age	68	Low	100%	-0.18	Yes

*t-test was used for categorical variables and Pearson or Spearman correlation was used for continuous variables as appropriate.

anxiety, depression, social participation, and sleep disturbance instruments, the pain disability index, and the patient global disease activity, and poorly with age (Table 3, Supplementary Table 1). PROMIS pain interference was similar between males and females (p=0.18).

The PROMIS fatigue instrument correlated strongly with the PROMIS physical function and the pain NRS, moderately with the PROMIS anxiety, depression, social participation, and sleep disturbance instruments, and the patient global disease activity (Table 3, Supplementary Table 1). PROMIS fatigue score was similar between males and females (p=0.65).

Floor/ceiling effects and internal consistency

The PROMIS physical function instrument did not demonstrate any significant ceiling (2%) or floor effect (9%). The PROMIS pain interference instrument had a significant ceiling effect. For interference instruments, maximal interference was interpreted as the worst possible score, or floor, and minimal interference as the best possible score, or ceiling. The PROMIS pain interference instrument was found to have poor sensitivity to detect low levels of pain (31% of patients reporting minimum pain interference). No floor effect (i.e., ability to detect high levels of pain) was observed (1% of patients reporting maximum pain interference score). There did not appear to be a significant ceiling (1%) or floor effect (<1%) for the PROMIS fatigue scale. For each instrument, internal consistency was high with a Cronbach α of 0.97 for pain interference, 0.89 for fatigue, and 0.96 for physical function.

Test-retest reliability

A total of 296 IIM participants initiated the survey (Table 2). Of the 296 participants, a variable number of participants completed both test and retest for each questionnaire (n=159 (Physical Function), 164 (Pain Interference), 162 (Fatigue)). For each respective questionnaire, 42%

(Physical Function), 41% (Pain Interference), and 41% (Fatigue) of participants did not complete the retest. Most participants who completed both test-retest were female (75%) with a mean age (SD) of 60.9 years (12.0). Participants who completed the baseline assessment only did not differ significantly in age or sex (p>0.05) (n=287, 59.6 (12.1), 76% female). The pain NRS was used as an anchor that was not expected to change and was unchanged between test-retest assessments (mean (SD), baseline 3.05 (2.6) vs retest 2.97 (2.55), p=0.76). Test-retest reliability was high for PROMIS physical function, pain and fatigue with ICC scores of (95% CI) 0.97 (0.96-0.98), 0.93 (0.91-0.95), 0.94 (0.91-0.95), respectively (Fig. 1, Table 2). Additionally, the mean difference between test and retest assessments for PROMIS physical function, pain interference, and fatigue was small and non-significant indicating no change.

Discussion

In this study, we report excellent internal consistency, high test-retest reliability, and strong construct validity of the PROMIS fatigue, pain interference and physical function instruments in a large cohort of adult patients with inflammatory myopathies. This is the first study assessing psychometric properties of the PROMIS physical function 8b, pain interference 6a, and fatigue 7a forms in adult patients with IIM. This study demonstrates that these PROMIS instruments may be appropriate tools for use in future myositis clinical studies for assessment of fatigue, pain, and physical function. Additional ongoing longitudinal validation studies are underway evaluating sensitivity to change and discrimination between groups of different levels of change.

For construct validity of the PROMIS pain interference and fatigue, 3 out of the 4 hypotheses were met for pain interference, and 4 out of 4 were met for fatigue supporting strong construct validity of these instruments. The only hypothesis that was not met for pain interference

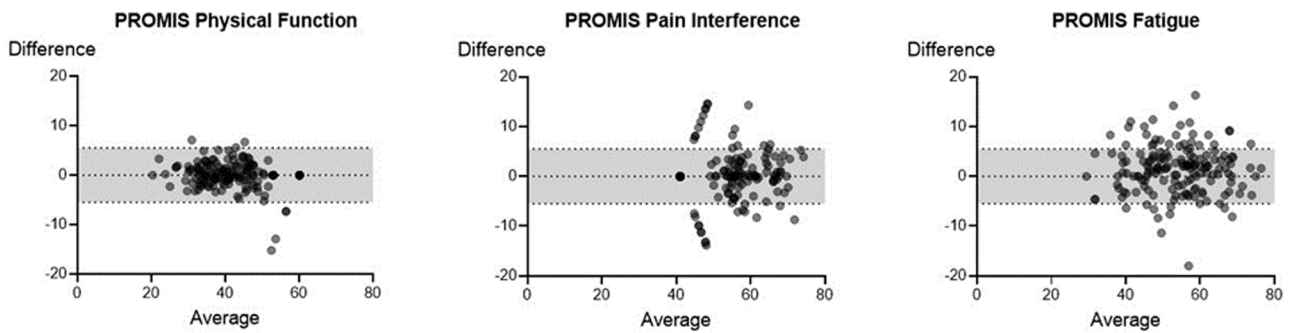


Fig. 1. Bland-Altman plots of the test-retest scores of the PROMIS physical function (n=158), pain interference (n=164), and fatigue (n=162) forms. Horizontal lines depict the mean and 95% confidence interval.

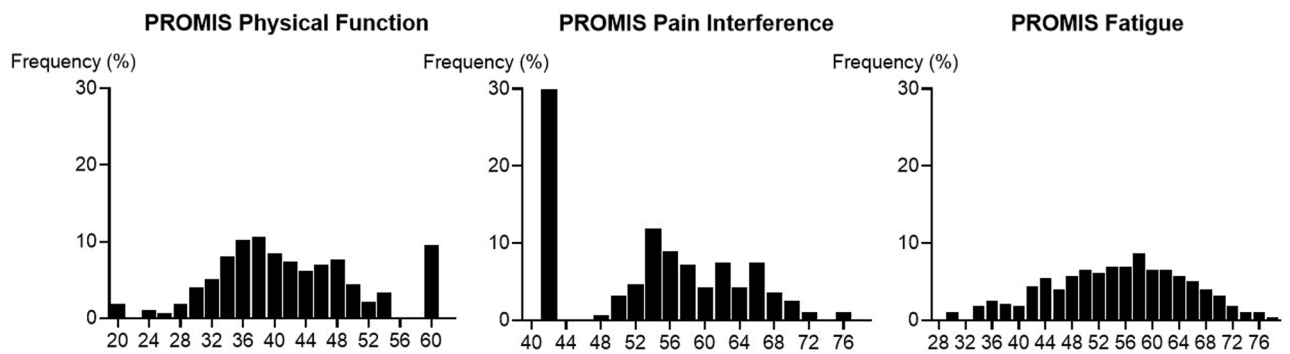


Fig. 2. Distribution of baseline T-Scores of the PROMIS* physical function (n=272), pain interference (n=279), and fatigue (n=275) forms. *For PROMIS physical function, pain interference, and fatigue forms, the higher scores indicate better physical function, more pain interference, and worse fatigue, respectively.

was degree of correlation between the PROMIS pain interference and pain disability index (PDI). Pain interference items assess the effect of pain on daily activities, while PDI indicates disability attributed to pain. Therefore, our *a priori* hypothesis was a high degree of correlation between these two instruments with higher pain levels associated with higher disability attributed to pain. However, the results showed a moderate correlation (0.63) between these tools instead of strong (>0.75) highlighting the complex relationship between pain and daily function. Another explanation could be the small number of patients with low levels of pain interference and pain related disability. Finally, if the correlation coefficient of $r=0.63$ was interpreted based on alternative prior literature, this may be considered as a strong correlation [30].

Besides strong construct validity, internal consistency was good for PROMIS fatigue ($\alpha=0.89$) and excellent for pain interference ($\alpha=0.97$), which raises concern for possible redundancy in the PROMIS pain interference tool. Previous OMERACT study with focus group interviews had signaled a similar concern by the attendants [14]. Especially, three items were thought to be potentially redundant according to previous work: “How much did pain interfere with your ability to participate in social activities?”, “How much did pain interfere with your enjoyment of social activities?”, and “How much did pain interfere with your household chores?”. Respondents had suggested to combine the first two questions into one question. However, given the pain interference questionnaire is already short with 6 questions, this is likely not a significant concern. Furthermore, although two of the pain interference questions related to ability to participate in social activities and enjoyment of those activities are similar, the PROMIS pain interference questionnaire was developed using Item Response Theory to cover the measurement continuum of the domain. The information function for these questions indicated that they provided unique information, allowing for higher levels of precision across the scale. Another concern about the PROMIS pain interference form is the substantial ceiling

effect, with approximately one third of the patients achieving the lowest score possible. The ceiling effect raises the question about the ability of the tool to detect differences or changes in the assessment of the patients with no or minimal pain. However, most participants with IIM did endorse some degree of pain interference in our study; therefore, the pain interference instrument is still a valuable tool in pain assessment. Using an instrument with more items, such as the PROMIS pain interference 8a, may help attenuate a ceiling effect but unfortunately this instrument was not available in all necessary languages.

PROMIS Physical Function-8b showed good test-retest reliability and moderate correlations with MAP supporting its construct validity. One of the *a priori* hypotheses for construct validity of Physical Function-8b was moderate correlations with physical activity measure, IPAQ, which was not met in this study. Physical activity is defined as bodily movement that results in energy expenditure, whereas physical function refers to ability to perform daily activities. The majority of routine daily activities are considered to be light to moderate intensity physical activities. However, those who have the ability to perform their daily activities do not necessarily have the ability to perform moderate-vigorous activities such as running or shoveling. Therefore, it is expected that physical function correlates with light intensity physical activity but may not correlate with moderate-vigorous physical activity. In a study with myositis patients wearing physical activity monitor (Actigraph®), physical activity variables of daily step, cadence, and vector magnitude correlated moderately with physical function measures of HAQ-DI, SF-36 physical function, and PROMIS physical function form 20 scores [3]. The strength of correlation between functional measures and physical activity was noted to decrease with increased intensities of activity. Similarly, in a study with patients with rheumatoid arthritis, HAQ scores were significantly associated with both IPAQ metabolic equivalent of task and accelerometry vector magnitude; but not significantly associated with vigorous activity per IPAQ and accelerometry [31]. In

summary, although our *a priori* hypothesis was not met between IPAQ and Physical Function-8b, this was not an unexpected result.

A previous study demonstrated adequate test-test reliability, construct validity, and responsiveness of PROMIS physical function 20 form in adult patients with IIM (6 PM, 24 DM, 9 NM and 11 with AS) [32]. Both Physical Function-20 and Physical Function-8b include questions about difficulty with walking, taking stairs, running errands, lifting groceries and performing household chores which requires both low and moderate-vigorous physical activities, whereas, Physical Function-20 also includes questions about self-care including dressing, bathing, washing, hair shampooing, and toilet use. Although Physical Function-20 is more comprehensive, it is lengthier (20 vs 8 questions) which may limit its use in the clinical setting; however, it is the same length as the HAQ-DI. Physical Function-20 showed good responsiveness to change in a longitudinal myositis cohort [32], while the OMERACT study assessing the responsiveness of Physical Function-8b is currently underway and will allow for comparison of responsiveness of these two physical function tools. Both tools have similarly excellent internal consistency with potential concern for redundancy; however, all the items in Physical Function-8 had gained strong support from the respondents in previous patient focus groups making redundancy less of a concern for Physical Function-8 [14]. Content validity of the Physical Function-20 has not yet been studied in myositis. Neither Physical Function-20 nor Physical Function-8 had significant floor or ceiling effects, which supports their ability to capture the full spectrum of physical function in patients with myositis.

A major strength of this study is the large cohort of participants with adult IIM. This is also the first study assessing reliability and validity of PROMIS physical function 8b, pain interference 6a, and fatigue 7a forms in IIM. Even though PROMIS measures have been extensively validated in other general and chronic disease populations, performing validation studies using these widely applicable measures is necessary to take unique disease characteristics into account and complies with the working guidelines set forth by OMERACT and the Food and Drug Administration [5]. Limitations of this study include a lack of information on serologic profile, disease subset, and disease activity of the enrolled participants. Furthermore, only 15% of the enrolled participants were from non-U.S. countries limiting the individual assessment of psychometric properties of these instruments for patients living outside of the U.S.

In conclusion, PROMIS physical function 8b, pain interference 6a, and fatigue 7a forms are PROMs which demonstrated strong construct validity and test-retest reliability in a large cohort of adult patients with IIM. Internal consistency of these instruments was excellent. Both PROMIS fatigue and Physical Function instruments did not show any ceiling or floor effect with some concern for significant ceiling effect in the pain interference instrument (recognizing that pain is not a universal feature of IIM). Further longitudinal studies to assess the responsiveness of these measures are currently ongoing in Sweden, Australia, South Korea, the Netherlands and the USA.

Funding

Portions of the work have been supported by the Rheumatic Diseases Research Core Center (P30-AR053503/AR070254) Core B from the National Institutes of Arthritis Musculoskeletal and Skin Diseases (NIAMS) of the National Institutes of Health (NIH). D.D. was supported through T32-AR048522. C.M. is supported by The Johns Hopkins Clinician Scientist Award, the Jerome L. Greene Foundation, and K23 AR075898-01. Dr. Christopher-Stine's work has been supported in part through the Huayi and Siuling Zhang Discovery Fund. Portions of the work have been supported by NuFactor, OptionCare, BioRx, Soleo Health, Walgreens, and US Bioservices. Dr. Alexanderson, Dr. Regardt and Dr. Lundberg are supported by the Swedish Rheumatism Association. Funding from OMERACT which receives unrestricted funds from over 23 clinical research and pharmaceutical companies.

Disclaimer

All statements in this report including its conclusions are solely those of the authors and do not reflect the opinions of the National Institutes of Health (NIH) or the National Institutes of Arthritis Musculoskeletal and Skin Diseases (NIAMS).

Declaration of Competing Interest

The authors have no conflicts of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.semarthrit.2022.152111.

References

- [1] Lundberg IE, Fujimoto M, Vencovsky J, Aggarwal R, Holmqvist M, Christopher-Stine L, et al. Idiopathic inflammatory myopathies. *Nat Rev Dis Primers* 2021;7(1):86.
- [2] Ponyi A, Borgulya G, Constantin T, Vancsa A, Gergely L, Danko K. Functional outcome and quality of life in adult patients with idiopathic inflammatory myositis. *Rheumatology (Oxford)* 2005;44(1):83–8.
- [3] Rockette-Wagner B, Saygin D, Moghadam-Kia S, Oddis C, Landon-Cardinal O, Allenbach Y, et al. Reliability, validity and responsiveness of physical activity monitors in patients with inflammatory myopathy. *Rheumatology (Oxford)* 2021.
- [4] Regardt M, Welin Henriksson E, Alexanderson H, Lundberg IE. Patients with polymyositis or dermatomyositis have reduced grip force and health-related quality of life in comparison with reference values: an observational study. *Rheumatology (Oxford)* 2011;50(3):578–85.
- [5] Administration FaD. Roadmap to patient-focused outcome measurement in clinical trials. [cited 11/28/2020]; Available from: .
- [6] Opinc AH, Brzezinska OE, Makowska JS. Disability in idiopathic inflammatory myopathies: questionnaire-based study. *Rheumatol Int* 2019;39(7):1213–20.
- [7] DiRenzo D, Bingham 3rd CO, Mecoli CA. Patient-reported outcomes in adult idiopathic inflammatory myopathies. *Curr Rheumatol Rep* 2019;21(11):62.
- [8] Maxwell LJ, Beaton DE, Boers M, D'Agostino MA, Conaghan PG, Grosskleg S, 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.
- [9] Alexanderson H, Del Grande M, Bingham 3rd CO, Orbai AM, Sarver C, Clegg-Smith K, et al. Patient-reported outcomes and adult patients' disease experience in the idiopathic inflammatory myopathies. report from the OMERACT 11 Myositis Special Interest Group. *J Rheumatol* 2014;41(3):581–92.
- [10] Regardt M, Basharat P, Christopher-Stine L, Sarver C, Bjorn A, Lundberg IE, et al. Patients' experience of myositis and further validation of a myositis-specific patient reported outcome measure - establishing core domains and expanding patient input on clinical assessment in myositis. Report from OMERACT 12 *J Rheumatol* 2015;42(12):2492–5.
- [11] Park JK, Mecoli CA, Alexanderson H, Regardt M, Christopher-Stine L, Casal-Dominguez M, 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.
- [12] Mecoli CA, Park JK, Alexanderson H, Regardt M, Needham M, de Groot I, et al. Perceptions of patients, caregivers, and healthcare providers of idiopathic inflammatory myopathies: an international OMERACT study. *J Rheumatol* 2019;46(1):106–11.
- [13] Regardt M, Mecoli CA, Park JK, de Groot I, Sarver C, Needham M, 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.
- [14] Esfandiary T, Park JK, Alexanderson H, Regardt M, Needham M, de Groot I, 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.
- [15] Allenbach Y, Mammen AL, Benveniste O, Stenzel W. Immune-mediated necrotizing myopathies working G. 224th ENMC international workshop: clinico-sero-pathological classification of immune-mediated necrotizing myopathies Zandvoort, The Netherlands, 14-16 October 2016. *Neuromuscul Disord* 2018;28(1):87–99.
- [16] Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med* 1975;292(7):344–7.
- [17] Lundberg IE, Tjarnlund A, Bottai M, Werth VP, Pilkington C, de Visser 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. *Arthritis Rheumatol* 2017;69(12):2271–82.
- [18] Mokkink LB, Prinsen CA, Bouter LM, Vet HC, Terwee CB. The COSMIN and how to select an outcome measurement instrument. *Braz J Phys Ther* 2016;20(2):105–13.
- [19] Handbook O. [cited March 11, 2021.]; Available from: <https://omeracthandbook.org/handbook>.

- [20] Bartlett SJ, Orbai AM, Duncan T, DeLeon E, Ruffing V, Clegg-Smith K, et al. Reliability and validity of selected PROMIS measures in people with rheumatoid arthritis. *PLoS One* 2015;10(9):e0138543.
- [21] Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol* 2010;63(11):1179–94.
- [22] PROMIS Health Measures. [cited 02/09/2022]; Available from: <https://www.healthmeasures.net/explore-measurement-systems/promis>.
- [23] Hewlett S, Dures E, Almeida C. Measures of fatigue: bristol rheumatoid arthritis fatigue multi-dimensional questionnaire (BRAFMQ), bristol rheumatoid arthritis fatigue numerical rating scales (BRAFNRS) for severity, effect, and coping, chandler fatigue questionnaire (CFQ), checklist individual strength (CIS20R and CIS8R), fatigue severity scale (FSS), functional assessment chronic illness therapy (Fatigue) (FACIT-F), multi-dimensional assessment of fatigue (MAF), multi-dimensional fatigue inventory (MFI), pediatric quality of life (PedsQL) multi-dimensional fatigue scale, profile of fatigue (ProF), short form 36 vitality subscale (SF-36 VT), and visual analog scales (VAS). *Arthritis Care Res (Hoboken)* 2011;63(Suppl 11):S263–86.
- [24] Tait RC, Pollard CA, Margolis RB, Duckro PN, Krause SJ. The pain disability index: psychometric and validity data. *Arch Phys Med Rehabil* 1987;68(7):438–41.
- [25] Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003;35(8):1381–95.
- [26] Alexanderson H, Lundberg IE, Stenstrom 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.
- [27] Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012;24(3):69–71.
- [28] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15(2):155–63.
- [29] Cronbach LJ, Shavelson RJ. My current thoughts on coefficient alpha and successor procedures. *Educ Psychol Measur* 2004;64(3):391–418.
- [30] Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med* 2018;18(3):91–3.
- [31] Hernandez-Hernandez V, Ferraz-Amaro I, Diaz-Gonzalez F. Influence of disease activity on the physical activity of rheumatoid arthritis patients. *Rheumatology (Oxford)* 2014;53(4):722–31.
- [32] Saygin D, Oddis CV, Dzanko S, Koontz D, Moghadam-Kia S, Ardalan K, et al. Utility of patient-reported outcomes measurement information system (PROMIS) physical function form in inflammatory myopathy. *Semin Arthritis Rheum* 2021;51(3):539–46.