

The PROMIS Fatigue_{FM} Profile: a self-report measure of fatigue for use in fibromyalgia

Anna L. Kratz¹ · Stephen Schilling² · Jenna Goesling³ · David A. Williams⁴

Accepted: 12 January 2016 / Published online: 28 January 2016
© Springer International Publishing Switzerland 2016

Abstract

Purpose Fibromyalgia (FM) is characterized by myriad symptoms and problems. Fatigue is one of the most common, distressing, and disabling symptoms in FM. The purpose of this study was to use fatigue item banks that were developed as part of the Patient-Reported Outcomes Measurement Information System (PROMIS) to devise a self-report measure of fatigue for use in individuals with FM.

Methods A sample of 532 adults with FM (age range = 18–77, 96.1 % female) completed the PROMIS

fatigue item bank. Factor analyses and item response theory analyses were used to identify dimensionality and optimally performing items. These data were used in combination with clinical input to select items for a fatigue self-report measure for use in FM.

Results Factor analyses revealed four distinct factors in the PROMIS fatigue item bank; items for each univariate subscale were identified by selecting four items with high item information values. A 16-item measure, the PROMIS Fatigue_{FM} Profile, consisting of four 4-item short forms reflecting fatigue experience (“intensity”) and fatigue impact in three subdomains—social, cognitive, and motivation—was created. The new PROMIS Fatigue_{FM} Profile short forms showed excellent internal reliability, low ceiling and floor effects, and equivalent or higher test information compared to the standard 4- and 7-item PROMIS fatigue short forms.

Conclusions The newly developed PROMIS Fatigue_{FM} Profile, a 16-item measure consisting of four 4-item short forms of self-reported fatigue severity, shows early evidence of good psychometric characteristics, provides the ability to use short forms that assess distinct aspects of fatigue experience and fatigue impact, and demonstrates equivalent or higher levels of test information compared to standard PROMIS fatigue short forms with similar number of items. The PROMIS Fatigue_{FM} Profile indicated fatigue experience and impact levels approximately 1.5 standard deviations above the normative sample mean across all short forms. Future work to evaluate the validity and reliability of this new measure in individuals with FM is needed.

Keywords Fibromyalgia · Fatigue · Self-report measure · Short form · PROMIS

Electronic supplementary material The online version of this article (doi:10.1007/s11136-016-1230-9) contains supplementary material, which is available to authorized users.

✉ Anna L. Kratz
alkratz@umich.edu; alkratz@med.umich.edu

Stephen Schilling
schillsg@isr.umich.edu

Jenna Goesling
jennago@med.umich.edu

David A. Williams
daveawms@umich.edu

¹ Department of Physical Medicine and Rehabilitation, University of Michigan, North Campus Research Complex, 2800 Plymouth Rd, Building NCRC B14, Room G218, Ann Arbor, MI 48109-2800, USA

² Institute for Social Research, University of Michigan, Ann Arbor, MI, USA

³ Department of Anesthesiology, University of Michigan, Ann Arbor, MI, USA

⁴ Departments of Anesthesiology, Medicine, Psychiatry, and Psychology, University of Michigan, Ann Arbor, MI, USA

Introduction

Although widespread pain is the cardinal symptom of fibromyalgia (FM), fatigue is one of the most common and impactful symptoms in FM [1–3]. The experience of fatigue in FM is described as profound and overwhelming, more severe, constant, and unpredictable than normal tiredness, not relieved by resting or sleep, not proportional to effort exerted, and disruptive in terms of motivation, activities, and cognition [4, 5]. Like other conditions characterized by multiple symptoms, rigorous assessment of all symptom domains in FM can be time-consuming and burdensome for respondents. Therefore, there is a need for valid, precise, and efficient measures of the health-related quality of life (HRQOL) domains that are of relevance to individuals with FM.

A primary goal of the NIH-funded initiative, the Patient-Reported Outcomes Measurement Information System (PROMIS), was to develop efficient, accessible, and high-quality measures of HRQOL that can be universally applied across conditions. A rigorous multi-step methodology that incorporated both classical test theory and item response theory (IRT) was used to develop PROMIS instruments. PROMIS measures are scored on a *T* score metric (Mean = 50, SD = 10) and are available as computer-adaptive tests (CAT), short forms (a static set of items), or profile measures (a fixed collection of short forms measuring different concepts). PROMIS currently includes a 95-item bank to assess fatigue in adults that can be administered as a CAT, standard 4-, 6-, 7-, or 8-item short forms, or as a customized short form. The PROMIS fatigue item bank contains items that assess the experience of fatigue (e.g., frequency, duration, intensity) and the impact of fatigue on physical, mental, and social activities. Additionally, some fatigue experience items assess fatigability, or fatigue severity in the context of demand/activity (e.g., “How often did you feel tired even when you hadn’t done anything?”).

Interestingly, despite the variability in item content, previous factor analyses of the PROMIS fatigue items suggest a single “general fatigue factor” and have supported the inclusion of all fatigue items in a single measure [6]. This is in contrast to the PROMIS tools to measure pain, which have separate measures of pain intensity (a 3-item short form) and pain interference (CATs and short forms) [7, 8]. This unidimensional approach to measuring fatigue is also inconsistent with a conceptual model of fatigue in FM indicating multiple facets of fatigue experience (e.g., overwhelming, weak/heavy feelings) and impact (e.g., motivational difficulties, problems thinking/concentrating) [5]. The authors of this conceptual model highlighted the difficulty in measuring fatigue in FM given that it is such a complex experience for most patients and suggested that a good patient-reported outcome

measure should assess all the characteristics of fatigue identified in the conceptual model [5].

Because the 95 PROMIS fatigue items provide a wide range of content coverage of the fatigue domain, this set of items has the potential to provide assessment of the various aspects of fatigue that are relevant to FM. Indeed, findings from a focus group study indicated that individuals with FM rated most PROMIS fatigue items as relevant to their condition, but most items reflected an underestimation of the profound experience of fatigue in FM [9]. These findings suggest the need for careful selection of items for measuring fatigue in FM. In a previous study, PROMIS fatigue items were selected by an expert panel to create a fatigue short form for use in multiple sclerosis [10], another condition where fatigue is a prominent and disabling symptom. The resulting PROMIS Fatigue_{MS} Short Form was found to have higher content validity than the standard PROMIS short form [10]. This study speaks to the potential to leverage the flexible design of the PROMIS to create customized condition-specific measures.

The aim of this study was to devise a fatigue self-report measure for use in FM by selecting a subset of existing PROMIS fatigue items. Item selection was guided by previous research and clinical experience relevant to fatigue in FM as well as factor analyses and IRT analyses to identify (1) the dimensions (i.e., factors) inherent in the PROMIS fatigue item bank in a sample of adults with FM and (2) the best performing items in adults with FM. The psychometric properties (reliability, distribution properties, test information) of the newly developed PROMIS Fatigue_{FM} Profile scales were examined and compared to four PROMIS standard fatigue short forms (i.e., the PROMIS Fatigue short forms 4a, 6a, 7a, and 8a).

Patients and methods

Participants

Participants were 532 adults who reported a diagnosis of FM and were members of the National Fibromyalgia Association (NFA), a patient advocacy organization. Data were collected from participants representing all 50 states of the USA. The majority of participants were female and Caucasian. Additional participant demographics are described in Table 1.

Study procedures

This study is based on data from the fatigue arm of a larger study ($N = 4265$) that examined multiple PROMIS item banks (e.g., pain interference, sleep disturbance) in persons

Table 1 Participants' descriptive data

(<i>N</i> = 532)	<i>N</i>	%	Mean	SD	Min–Max
Age (years)			49.23	10.98	18–77
Time since FM diagnosis (years)			11.92	9.94	0.5–53.0
Sex (female)	497	96.1			
Race					
White	346	65.0			
Black/African Am	5	0.9			
American Indian/Alaska Native	1	0.2			
Asian	3	0.6			
Pacific Islander	4	0.8			
Other/not provided	172	32.3			
Education					
Eighth grade	1	0.2			
Some high school	9	1.7			
High school/GED	80	15.0			
Some college/technical degree	242	45.5			
College degree (BA/BS)	120	22.6			
Advanced degree (MA, PhD, MD)	65	12.2			
Other/not provided	15	2.8			

Participants were not required to complete all demographic items; amount of missing data varies across the variables. *GED* General Educational Development

with FM. So that no participant had to complete all PROMIS item banks, which would be excessively burdensome, each participant only completed 1 bank of items plus a few calibration items from other item banks. The online data collection platform was programmed so that once a participant completed the demographic portion of the survey, they were randomly assigned to complete one of the item banks. The subsample of volunteers who were assigned to the fatigue arm were not different from the full sample in terms of distribution of sex ($X^2 = 0.02$, $p = 0.89$), race ($X^2 = 18.93$, $p = 0.06$), marital status ($X^2 = 4.93$, $p = 0.42$), education level ($X^2 = 0.94$, $p = 0.99$), or average age [mean difference = 0.07 years; $F(1, 3876) = 0.019$, $p = 0.89$]. The Institutional Review Board at the University of Michigan approved the study before it was initiated. Participants were recruited through an advertisement in the monthly NFA newsletter that was distributed to approximately 70,000 NFA members. The advertisement described the study and provided a URL (Web site) link to the study site (Assessment CenterSM). Those who entered the study Web site first viewed an informed consent page, and volunteers could indicate consent to participate in the survey study electronically. Once they had consented and were enrolled/registered in the study, participants were sent a unique Assessment CenterSM login ID and password. Participants were

compensated \$10. Data for this study were collected between April 2009 and May 2010.

Measures

After completing a basic demographic form, each participant completed the 95-item PROMIS fatigue item bank in addition to a small subset of items from other item banks (e.g., depression, anxiety).

Data analyses

Exploratory and confirmatory factor analyses

Full-information item factor analysis [11] using a computer program for conducting ordinal factor analysis, ORDFAC [12], was used to examine factor structure. This approach produces results that are slightly more accurate and that allow for greater clarity in the identification of the number of factors compared to factor analysis of polychoric correlation matrices [11]. Although both the Akaike information criterion (AIC [13]) and the Bayesian information criterion (BIC [14]) were used to compare model fit to the data, emphasis was placed on BIC values, given that this is the more conservative of the two criteria. For both exploratory and confirmatory factor analyses, the model with the lowest AIC and BIC values was chosen.

Item response theory analyses

Using the same modeling approach as has been used by PROMIS to calibrate item banks, Samejima's graded response model (GRM) [15] was used to fit the item responses to each unidimensional fatigue subscale that was identified through factor analysis. GRM fit includes defining the slope parameter, which details the steepness of the item characteristic curves (ICC's) and the threshold parameters, which gives the intersections of the ICC's for adjacent categories. All modeling was completed with the MULTILOG 7 [16] statistical software. The R statistical package [17] was used to estimate the item/test polyserial correlations.

Results

Factor analysis

Exploratory factor analyses

For the exploratory analysis, the BIC indicated a five-factor model (see Table 2). The PROMAX-rotated factor loadings presented an easily interpretable solution and are presented in Online Resource 1. The first factor contained negatively worded fatigue experience items (e.g., tired, fatigued, exhausted). The factor loadings showed that the fifth factor consisted solely of fatigue experience items that were positively worded and assessed how energetic the respondent felt. It is not uncommon in a large item bank such as this for items that assess the same construct to be split based on positive/negative wording of the item (e.g., the PROMIS cognitive functioning item banks [18]). Therefore, it is reasonable to consider the negatively worded items on factor 1 and the positively worded items on factor 5 to represent a single "fatigue experience" dimension.

Items that assess fatigue impact were split across three separate factors: cognitive impact (factor 2; e.g., memory,

attention, decision making), social impact (factor 3; e.g., socializing, recreational activities), and motivational impact (factor 4; starting or finishing things, pushing oneself to get things done). Distinguishing between different types of fatigue, including physical fatigue, motivational fatigue, cognitive fatigue, and reduced activity is consistent with other legacy measures of fatigue, such as the Multidimensional Fatigue Inventory (MFI [19]) which has been used extensively in clinical trials of FM [20].

Confirmatory factor analysis

Results from a confirmatory factor analysis (Table 3) support the conclusion that a 4-factor solution is the best fit to these data and is a better fit than either a 2-factor or 1-factor solution. Notably, the 2-factor solution was significantly better than the 1-factor solution. Taken together, these factor analyses provide evidence that the PROMIS fatigue items should not be considered a unidimensional construct for individuals with FM, as they have been in other samples [6].

IRT results

The primary aim of the IRT analyses was to identify the items that were most informative for measuring each dimension of fatigue in this sample of individuals with FM. Item information was determined by steeper slopes that indicate higher item information functions (IIF's). The most informative items were evaluated relative to other high information items that represented a broad range of the fatigue spectrum. This approach was undertaken to maximize accessibility/usability along with high IIF's.

Creation of a 16-item PROMIS Fatigue_{FM} Profile

Because four fatigue factors were identified, the aim was to devise a profile measure consisting of four short forms, one for each unidimensional factor. Items for the PROMIS Fatigue_{FM} Profile were selected based on a number of criteria. It was important to select items that demonstrated excellent, or "high discrimination" IRT parameters; because each profile short form was designed to have as few items as possible while still maintaining adequate levels of reliability, it was important that each item provides a great deal of information. Four items were selected for each short form—fatigue experience, fatigue cognitive impact, fatigue social impact, and fatigue motivational impact. To aid the study team in item selection, items within each factor were listed in rank order from highest to lowest average item information. Items were considered for selection starting with the highest information item and working down the list. Items were skipped/not included if

Table 2 Fatigue item exploratory factor analysis model comparisons ($N = 532$)

Model	Log-Lik	Chi-Sq	DF	AIC	BIC
1 Factor	-45,092	90,185	95	90,375	90,781
2 Factor	-43,340	86,680	189	87,058	87,865
3 Factor	-42,074	84,148	282	84,712	85,917
4 Factor	-41,493	82,986	374	83,734	85,332
5 Factor	-41,195	82,390	465	83,320	85,306

Log-Lik log likelihood, *Chi-Sq* Chi-squared, *AIC* Akaike information criterion, *BIC* Bayesian information criterion

Table 3 Fatigue item confirmatory factor analysis model comparisons ($N = 532$)

Model	NPARAM	Log-Lik	Chi-Sq	Chi-Diff	p value	AIC	BIC
1 Factor	465	-45122.6	90245.3			91175.3	93162.2
2 Factor	466	-43853.0	87706.0	2539.3	1.0E-207	88638.0	90629.2
4 Factor	471	-43819.0	87638.1	67.9	2.8E-13	88580.1	90592.6

NPARAM number of independent parameters, *Log-Lik* log likelihood, *Chi-Sq* Chi-squared, *Chi-Diff* Chi-squared difference, *AIC* Akaike information criterion, *BIC* Bayesian information criterion

they were highly redundant with already selected items (e.g., “To what degree did you have trouble starting things because of fatigue?”, “How often did you have trouble starting things because of fatigue?”) in order to maximize content coverage within each dimension. We also made an effort to select items that did not contain colloquial wording (e.g., “wiped out,” “run down,” “physically drained”) for more straightforward translation to other languages and to optimize comprehension of the scale for a broader range of English speakers.

Four is the minimum number of items per short form in current PROMIS profile measures; therefore, it was decided to select four items per short form. However, the reliability and test information of short forms with 3 items were also examined to determine whether a shorter measure could perform as well as a longer measure. Short forms consisting of the three items with the highest average item information values resulted in moderate reductions in test reliability statistics (mean change = -0.04) and lower test information, such that all but the fatigue social impact short form demonstrated lower test information compared to the PROMIS Fatigue Short Form 4a. Because an additional item per short form represents negligible added respondent burden with the benefit of greater reliability and test information, it was determined that the four item short forms should be retained.

Overview of the 16-item PROMIS Fatigue_{FM} Profile

The items and response options for the four short forms of the PROMIS Fatigue_{FM} Profile are found in Table 4. Four of the items reflect intensity of fatigue and constitute the fatigue experience short form, using the words fatigue/fatigued, tired, and exhausted. The fatigue social impact short form contains four items that assess the degree to which fatigue interferes with or limits social activities, recreational activities, and socializing with friends. The fatigue motivational impact short form contains items that assess the degree to which fatigue interferes with starting or finishing things, effectiveness at home, and having to push oneself to get things done. Finally, the fatigue cognitive impact short form contains four items that cover the impact of fatigue on speed and clarity of thought, trouble making decision, or remembering. All items are negatively

worded and are scored in the same direction; each short form is scored by summing the items. The possible raw score range is 4–20 for each short form, with higher scores indicating greater fatigue experience/impact.

In order to be comparable with the PROMIS scoring metric, two sets of T score scale conversions were calculated for the PROMIS Fatigue_{FM} Profile experience and impact short forms (Table 5). The study sample T score conversion values are based on IRT parameters from this study; consequently, T scores obtained from this conversion would provide for comparison with this sample of 532 people with self-reported FM. The PROMIS-equivalent T score conversion values are based on IRT parameters from PROMIS; thus, T scores derived from this conversion provide for comparison with the PROMIS normative sample. This allows the user to make comparisons with this sample (an approximation of the population of interest), or ostensibly with both this sample and the normative population. Notably, a raw score on the PROMIS Fatigue_{FM} Profile or short forms would result in a much lower T score on the study-equivalent T score scale compared to the PROMIS-equivalent T score scale. This is due to the fact that the sample of individuals with FM has higher levels of fatigue compared to the general populations that is represented by the PROMIS normative sample.

Pearson correlations between the PROMIS Fatigue_{FM} Profile short forms (see Table 6) range from $r = 0.60$ to 0.77 (i.e., 36–59 % shared variance), suggest a relatively modest degree of overlap between the unidimensional subscales, and support the creation of short forms within the profile measure.

Reliability, data distribution, and test information results

Reliability and data distribution results are depicted in Table 7. The PROMIS Fatigue_{FM} Profile short forms and the standard PROMIS short forms all demonstrated good-to-excellent internal consistency reliability; the PROMIS Fatigue Short Form 7a demonstrated the lowest reliability. Notably, T score (PROMIS metric) means across the scales indicate high levels of fatigue in this sample. There is a high degree of consistency across the scales, indicating that this sample reports average fatigue levels 1.5 standard deviations above the normative sample mean of 50; the one

Table 4 PROMIS Fatigue_{FM} Profile

No.	Item Name	Short Form	In the past 7 days...	Not at all	A little bit	Somewhat	Quite a bit	Very much
E1	FATEXP34	Experience	How tired did you feel on average?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
E2	FATEXP40	Experience	How fatigued were you on average?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
E3	FATEXP36	Experience	How exhausted were you on average?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
S1	FATIMP48	Social Impact	To what degree did your fatigue interfere with your social activities?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
S2	FATIMP45	Social Impact	To what degree did your fatigue interfere with your ability to engage in recreational activities?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
M1	FATIMP27	Motivational Impact	To what degree did you have trouble starting things because of your fatigue?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
M2	FATIMP51	Motivational Impact	To what degree did you have trouble finishing things because of your fatigue?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
C1	FATIMP38	Cognitive Impact	To what degree did your fatigue make it difficult to make decisions?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
C2	FATIMP2	Cognitive Impact	To what degree did your fatigue make you feel slowed down in your thinking?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
No.	Item Name	Short Form	In the past 7 days...	Never	Rarely	Sometimes	Often	Always
M3	FATIMP42	Motivational Impact	How often were you less effective at home due to your fatigue?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
M4	FATIMP3	Motivational Impact	How often did you have to push yourself to get things done because of your fatigue?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
S3	FATIMP18	Social Impact	How often did you have to limit your social activities because of your fatigue?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
S4	FATIMP56	Social Impact	How often were you too tired to socialize with your friends?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
C3	FATIMP30	Cognitive Impact	How often were you too tired to think clearly?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
C4	FATIMP11	Cognitive Impact	How often did your fatigue make you more forgetful?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5
			In the past 7 days...	None	Mild	Moderate	Severe	Very Severe
E4	FATEXP56	Experience	What was the level of your fatigue on most days?	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5

Item name indicates the name used to identify the item within the PROMIS item bank, currently available at www.assessmentcenter.net

exception is the PROMIS Fatigue Short Form 7a which showed a lower average fatigue level, less than 1 standard deviation above the PROMIS normative sample mean.

None of the measures showed problems with floor effects (proportion with lowest possible score) or ceiling effects (proportion with highest possible score). The PROMIS

Table 5 PROMIS Fatigue_{FM} Profile subscale raw scores (sum of all item scores) and equivalent *T* scores for this study sample and for the PROMIS normative sample

Fatigue experience raw score	Study sample <i>T</i> score	PROMIS [®] -equivalent <i>T</i> score	Fatigue social impact raw scores	Study sample <i>T</i> score	PROMIS [®] -equivalent <i>T</i> score
4	19	33	4	21	38
5	24	39	5	25	45
6	27	42	6	27	48
7	30	45	7	30	50
8	33	48	8	32	52
9	36	51	9	34	54
10	39	53	10	36	56
11	42	55	11	38	57
12	45	57	12	41	59
13	48	59	13	43	61
14	51	62	14	46	63
15	55	64	15	48	64
16	58	67	16	51	66
17	61	70	17	54	68
18	67	71	18	57	70
19	67	74	19	60	73
20	67	78	20	66	77
Fatigue motivational impact raw score	Study sample <i>T</i> score	PROMIS [®] -equivalent <i>T</i> score	Fatigue cognitive impact raw score	Study sample <i>T</i> score	PROMIS [®] -equivalent <i>T</i> score
4	13	36	4	20	39
5	17	42	5	24	45
6	21	45	6	28	49
7	24	47	7	31	51
8	27	49	8	34	54
9	30	52	9	36	56
10	33	54	10	39	58
11	36	56	11	41	60
12	38	58	12	43	61
13	41	60	13	46	63
14	44	61	14	48	65
15	47	63	15	51	67
16	50	65	16	54	69
17	54	68	17	57	71
18	57	70	18	60	74
19	61	73	19	64	77
20	66	77	20	69	81

Fatigue Short Form 7a, which was specifically designed to cover the extreme tails of the fatigue distribution, had zero cases at the floor or ceiling, whereas the other measures demonstrated very low floor/ceiling rates. Scores in this sample for all PROMIS measures approximated a normal distribution, as evidenced by low skew and kurtosis values.

Test information curves (Fig. 1) indicate that all of the short forms of the PROMIS Fatigue_{FM} Profile provided

generally equivalent or greater test information than standard 4- or 7-item PROMIS Fatigue short forms in the range of fatigue severity between -2.0 and $+2.0$ standard deviations. Notably, two of the PROMIS Fatigue_{FM} Profile short forms, fatigue cognitive impact and fatigue social impact, perform as well or better than all standard PROMIS short forms, even those with 6 or 8 items, across the fatigue spectrum. For fatigue experience and fatigue motivational

Table 6 Bivariate Pearson correlations between PROMIS Fatigue_{FM} Profile short form scores

	2	3	4
1. Fatigue experience	0.70	0.75	0.60
2. Fatigue social impact	–	0.77	0.62
3. Fatigue motivational impact		–	0.70
4. Fatigue cognitive impact			–

impact, the standard 6- or 8-item PROMIS short forms provided more information, but this is expected given they contain more items. The PROMIS Fatigue short forms 6a and 8a provide slightly more information than the newly created short forms at the lowest levels of fatigue (<-2.0 standard deviations). The Fatigue Short Form 7a provides slightly more information at the highest levels of fatigue ($>+2.0$ standard deviations), but over the entire range of fatigue, the 7-item form provided less information compared to either the new short forms or the more generic PROMIS Fatigue short forms 6a and 8a.

Discussion

The federally funded initiative, PROMIS, has devised a large battery of item banks that are meant to be universally applicable across chronic conditions [7, 21]. However, examination and validation of PROMIS measures within a given clinical population is recommended before using them within that group [21, 22]. The PROMIS design allows for customized measures that are created by selecting a subset of items that are deemed optimal for a given clinical population. Previous work has capitalized on this flexible feature of PROMIS to develop condition-specific short forms for fatigue in MS [10] and cognitive dysfunction in FM [23]. This study utilized a combination of classical test theory and item response theory to develop an FM-specific measure of fatigue and to compare the new

measure to existing standard PROMIS short forms of the same construct.

The 16-item PROMIS Fatigue_{FM} Profile development is consistent with the PROMIS approach to profile measurement development, which involves devising a static set of short forms, each of which consists of the best items for a given domain or subdomain. As with the PROMIS standard profile measures, the PROMIS Fatigue_{FM} Profile provides raw and *T* scores for each short form with no overall (total) score, due to the fact that the short forms together do not represent a unidimensional construct. It was developed, guided by knowledge of the experience of fatigue in FM (based on previous research and the clinical experience of three of the authors) as well as psychometric indicators, including factor analysis and item information function results. One strength of the PROMIS Fatigue_{FM} Profile is that it has four unidimensional short forms that capture distinct aspects of fatigue in FM. The fatigue experience short form contains items that assess intensity of tiredness, exhaustion, and fatigue. The three fatigue impact short forms contain items that assess the impact of a person's fatigue on his or her mental, social, and motivational activities and abilities. The capability to measure distinct facets of fatigue might be particularly important when using the PROMIS Fatigue_{FM} Profile as a treatment outcome measure, where it is expected that the intervention will have relatively stronger impact on one of subdomains of fatigue in FM. In contrast, the PROMIS Fatigue Short Form 8a, which includes 4 experience and 4 impact items, the PROMIS Fatigue Short Form 7a, which has 3 experience and 4 impact items, the PROMIS Fatigue Short Form 6a, which contains 3 impact and 3 experience items, and the PROMIS Fatigue Short Form 4a, which contains 3 experience and 1 impact item, were all designed to be scored and used only as a total scale due to the fact that factor analyses in the normative population revealed a single fatigue factor. Interestingly, the correlation between subscales that assess fatigue experience and fatigue impact was substantially smaller in

Table 7 Reliability and data distribution characteristics of the fatigue PROMIS Fatigue_{FM} Profile short form scores and standard PROMIS Fatigue short forms

Measure	Reliability	<i>T</i> score Mean \pm SD	Floor (%)	Ceiling (%)	Skew	Kurtosis
PROMIS Fatigue _{FM} Profile short forms						
Fatigue experience	0.89	65.39 \pm 6.86	0	7.1	–0.38	–0.20
Fatigue social impact	0.91	65.16 \pm 6.89	0.6	10.2	–0.69	0.22
Fatigue motivational impact	0.88	65.16 \pm 6.29	0	8.3	–0.63	0.34
Fatigue cognitive impact	0.90	65.91 \pm 7.21	0.6	4.9	–0.48	–0.02
PROMIS Fatigue Short Form 4a	0.86	64.83 \pm 6.31	0	11.1	–0.57	0.02
PROMIS Fatigue Short Form 6a	0.90	65.30 \pm 6.16	0	7.9	–0.67	0.20
PROMIS Fatigue Short Form 7a	0.83	59.85 \pm 5.33	0	0	–0.47	–0.68
PROMIS Fatigue Short Form 8a	0.87	65.51 \pm 6.04	0	5.6	–0.68	0.41

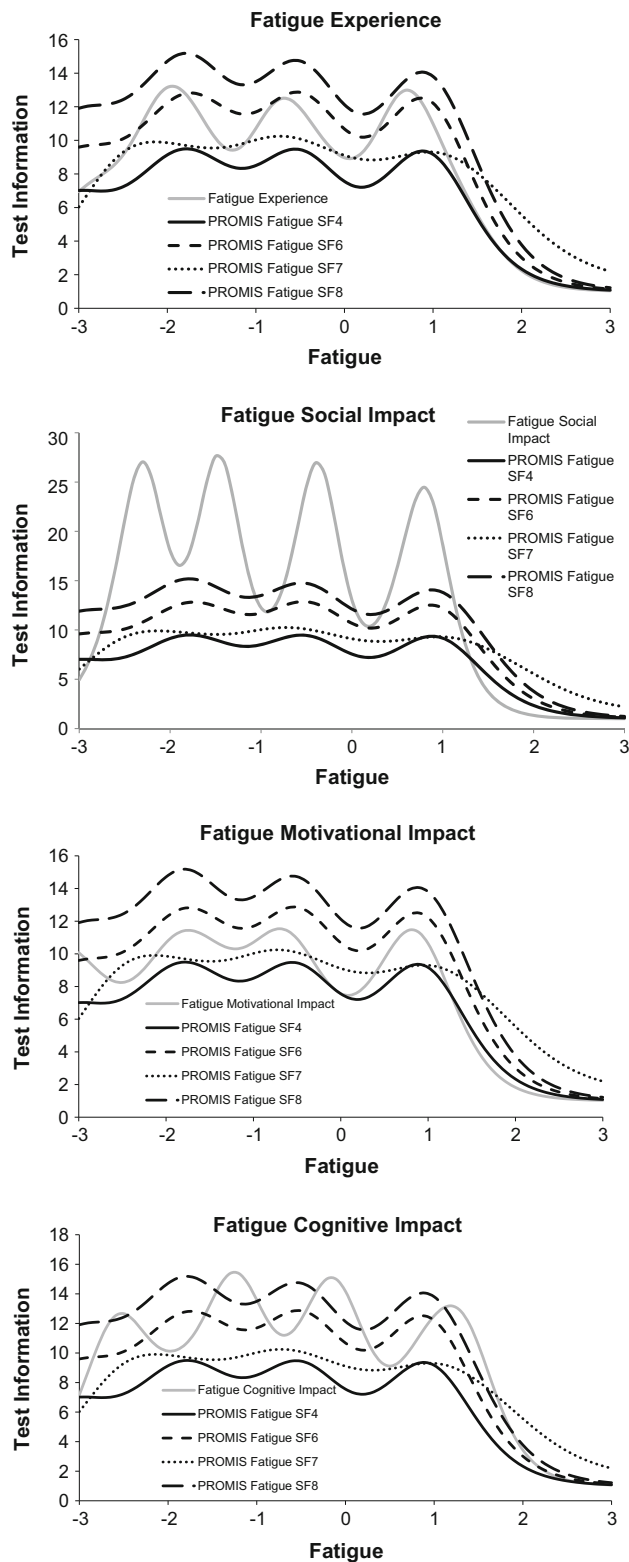


Fig. 1 Test information curves comparing PROMIS Fatigue_{FM} Profile short forms—fatigue experience, fatigue social impact, fatigue motivational impact, and fatigue cognitive impact—and PROMIS Fatigue short forms 4a, 6a, 7a, and 8a. The fatigue *x*-axis depicts the FM sample theta (SD) scale

this study (r 's = 0.60–0.75; 36–56 % shared variance) compared to the PROMIS development sample ($r = 0.95$; 90 % shared variance) [6]. The very high correlation in the development sample was interpreted as evidence that the items that assess fatigue experience and fatigue impact could be included in a single unidimensional measure. The results from this study suggest that dimensionality of the item bank may differ across samples and different clinical conditions and highlights the need for examination of the psychometric properties of PROMIS measures in specific populations before they can be confidently adopted for use.

The PROMIS Fatigue_{FM} Profile performed well psychometrically in this sample. Reliability was excellent for each of the short forms (experience = 0.89; social impact = 0.91; motivational impact = 0.88; cognitive impact = 0.90). Because we used the same sample to select the best performing items and to evaluate the psychometrics of the resulting subscales, it is possible that reliability calculations reflect as overestimation of what would be found in other samples of adults with FM. Reliability of these scales should be re-evaluated in other FM samples. Further examination of the psychometric properties, including construct validity, test–retest reliability, sensitivity to change, and responsiveness/minimally important differences, of the PROMIS Fatigue_{FM} Profile is also needed. Additional comparisons of the psychometric performance of the PROMIS Fatigue_{FM} Profile with standard PROMIS short forms and CATs in FM are warranted. However, it is recommended that clinicians and/or researchers using PROMIS to assess fatigue in FM use the PROMIS Fatigue_{FM} Profile. Although the PROMIS Fatigue short forms 6a and 8a also performed satisfactorily in this sample, the use of the newly developed PROMIS Fatigue_{FM} Profile is encouraged for use in people with FM for a number of reasons. Item selection for the PROMIS Fatigue_{FM} Profile was based, in part, on factor analyses and IRT that indicated distinct domains of functioning that were important to assess in FM and items that offered the most information in assessing those domains. The PROMIS Fatigue Short Form 8a, the best performing standard PROMIS short form, contains four fatigue impact items that, unfortunately, only represent the motivational impact factor identified in this study; it does not contain items that assess the social or cognitive impact of fatigue. The PROMIS Fatigue_{FM} Profile, in contrast, contains subscales that reflect all of the fatigue domains that were revealed in factor analyses and are known to be important in the lives of people with FM [2]. The PROMIS Fatigue_{FM} Profile can be thought of as containing PROMIS fatigue item bank items that have been essentially “vetted” and selected for use in FM. Uniform use of the PROMIS Fatigue_{FM} Profile by researchers will allow for comparison of fatigue

experience and fatigue impact scores across studies. In addition, because it is possible to convert the PROMIS Fatigue_{FM} Profile scores to *T* scores, clinicians and researchers can compare the score for any given patient or study sample to others with FM (using this study sample *T* scores) and/or the general US population (using the PROMIS *T* scores).

The PROMIS Fatigue Short Form 7a is not recommended for use in FM due to its low test information and unusually low average fatigue score compared to other PROMIS measures. The items for the standard 4-, 6-, and 8-item fatigue short form were ranked and selected based on two empirical criteria: (1) maximum interval information and (2) CAT simulations using 1000 cases. In contrast, items for the 7-item fatigue short form were selected by an expert panel who were focused on item content and representing the full range of the fatigue construct. This approach resulted in a measure that contains more “high-difficulty” items (e.g., “how often did you experience extreme exhaustion?”, “how often were you too tired to take a bath or shower?”) and has more information at the highest (most severe) end of the fatigue spectrum compared to any of the other short forms being considered in this study. Although these high-difficulty items result in no cases at the ceiling of the 7-item measure in this sample, the test information across the range of the fatigue spectrum was relatively low, especially for a measure with a relatively high number of items. Although the PROMIS Fatigue Short Form 4a had adequate reliability in this sample, it had lower test information than any of the unidimensional short forms from the PROMIS Fatigue_{FM} Profile.

Study strengths and limitations

Strengths of this study include the large and geographically diverse sample, use of both classical test theory and IRT, and the balanced use of empirical information along with clinical and pragmatic consideration in the selecting of short form items. The study was limited by self-report of FM diagnosis; however, this is consistent with the methodology that was used in the development of the PROMIS measures. Study participants were NFA members, which makes it more likely that the self-reported diagnostic status was accurate; furthermore, 98.5 % of participants responded “yes” to the questions, “Have you been diagnosed with FM?” However, the recruitment of participations through the NFA alone might have resulted in a study sample that is not wholly representative of the FM population.

The limited types of data collected from participants in this study (e.g., basic demographics, PROMIS items) make it difficult to compare this sample with the broader FM

population; however, the sociodemographic characteristics of this study sample are similar to those for people with FM in general [3]. And in a similar study where we have collected online survey data from people recruited through the NFA, the sample was very comparable to other FM samples in terms of self-reported cognitive functioning and FM severity and represented a broad range of FM severity [23]. Because the sample lacked heterogeneity in terms of sex and race, our ability to generalize our findings to males or racial minorities is limited. Future research examining the psychometric performance of the measure needs to include more diverse samples, particularly to examine whether psychometric properties are consistent or different (e.g., differential item functioning) across subgroups of people with FM. Although it was important to provide sample-specific *T* scores so that users can see how an individual patient or study participant compares to a large sample of individuals with FM, future research in larger, well-characterized samples may result in different *T* score conversion values; such research is needed for the continued clinical and research utility of the PROMIS Fatigue_{FM} Profile.

Conclusion

In summary, fatigue is one of the hallmark symptoms of FM and it is important to assess fatigue clinically and for research purposes. Previous factor analyses of the PROMIS fatigue items suggested that fatigue could be measured as a unidimensional construct. However, the results from this study provide evidence that the fatigue PROMIS item bank has multiple dimensions in a sample of individuals with FM. This study leveraged the flexibility of the PROMIS items banks to develop a customized PROMIS Fatigue_{FM} Profile that contains four short forms, fatigue experience, fatigue social impact, fatigue motivational impact, and fatigue cognitive impact, which reflect the different dimensions identified in the factor analyses. This newly developed PROMIS Fatigue_{FM} Profile provides a broader assessment of fatigue in FM with overall better psychometric qualities than the standard PROMIS short forms. The PROMIS Fatigue_{FM} Profile provides clinicians and researchers a comprehensive yet relatively brief measure of fatigue that is psychometrically sound and is recommended for use by those who want to use a PROMIS measure of fatigue in FM.

Acknowledgments Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases under award No. U01AR55069-01 (PI: Williams) and award No. 1K01AR064275 (PI: Kratz). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interest to report.

Ethical approval This study was approved by the relevant research ethics committee.

Informed consent Informed consent was obtained from all participants in the study.

References

- Mease, P. J., Arnold, L. M., Crofford, L. J., Williams, D. A., Russell, I. J., Humphrey, L., et al. (2008). Identifying the clinical domains of fibromyalgia: Contributions from clinician and patient Delphi exercises. *Arthritis and Rheumatism*, *59*(7), 952–960.
- Bennett, R. M., Jones, J., Turk, D. C., Russell, I. J., & Matallana, L. (2007). An internet survey of 2,596 people with fibromyalgia. *BMC Musculoskeletal Disorders*, *8*, 27.
- Wolfe, F., Ross, K., Anderson, J., Russell, I. J., & Hebert, L. (1995). The prevalence and characteristics of fibromyalgia in the general population. *Arthritis and Rheumatism*, *38*(1), 19–28.
- Arnold, L. M., Bradley, L. A., Clauw, D. J., Glass, J. M., & Goldenberg, D. L. (2008). Evaluating and diagnosing fibromyalgia and comorbid psychiatric disorders. *Journal of Clinical Psychiatry*, *69*(10), e28.
- Humphrey, L., Arbuckle, R., Mease, P., Williams, D. A., Samsøe, B. D., & Gilbert, C. (2010). Fatigue in fibromyalgia: a conceptual model informed by patient interviews. *BMC Musculoskeletal Disorders*, *11*, 216.
- Lai, J. S., Cella, D., Choi, S., Junghaenel, D. U., Christodoulou, C., Gershon, R., & Stone, A. (2011). How item banks and their application can influence measurement practice in rehabilitation medicine: A PROMIS fatigue item bank example. *Archives of Physical Medicine and Rehabilitation*, *92*(10 Suppl), S20–S27.
- Cella, D., Riley, W., Stone, A., Rothrock, N., Reeve, B., Yount, S., et al. (2010). The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology*, *63*(11), 1179–1194.
- Amtmann, D., Cook, K. F., Jensen, M. P., Chen, W. H., Choi, S., Revicki, D., et al. (2010). Development of a PROMIS item bank to measure pain interference. *Pain*, *150*(1), 173–182.
- Williams, D. A., Schilling, S., Shibata, K. N., Zwinck, L., Matallana, L., & Arnold, L. (2009). Relevance of PROMIS item banks to individuals with fibromyalgia. *Arthritis and Rheumatism*, *60*(10), S1404.
- Cook, K. F., Bamer, A. M., Roddey, T. S., Kraft, G. H., Kim, J., & Amtmann, D. (2012). A PROMIS fatigue short form for use by individuals who have multiple sclerosis. *Quality of Life Research*, *21*(6), 1021–1030.
- Schilling, S. G., & Bock, R. D. (2005). High-dimensional maximum marginal likelihood item factor analysis by adaptive quadrature. *Psychometrika*, *70*(3), 1–23.
- Schilling, S. G. (2005). *ORDFAC: A computer program for ordinal factor analysis*. Ann Arbor: University of Michigan.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.
- Schwartz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464.
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph*, *17*, 1–100.
- Thissen, D., Chen, W. H., & Bock, R. D. (2003). *MULTILOG7: Multiple category item analysis and test scoring using item response theory*. Chicago, IL: Scientific Software International Inc.
- Ihaka, R. (1988). *R: Past and future history*. Auckland, New Zealand: The University of Auckland.
- Lai, J. S., Butt, Z., Wagner, L., Sweet, J. J., Beaumont, J. L., Vardy, J., et al. (2009). Evaluating the dimensionality of perceived cognitive function. *Journal of Pain and Symptom Management*, *37*(6), 982–995.
- Smets, E. M., Garssen, B., Bonke, B., & De Haes, J. C. (1995). The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *Journal of Psychosomatic Research*, *39*(3), 315–325.
- Arnold, L. M., Williams, D. A., Hudson, J. I., Martin, S. A., Clauw, D. J., & Crofford, L. J., et al. (2011). Development of responder definitions for fibromyalgia clinical trials. *Arthritis and Rheumatism*, *64*(3), 885–894.
- Cook, K., Kallen, M., Cella, D., Crane, P., Eldadah, B., & Hays, R., et al. (2014). The patient reported outcomes measurement information system (PROMIS) perspective on: Universally-relevant vs. disease-attributed scales. http://www.nihpromis.org/Documents/Universally-Relevant_vs_Disease-Attributed_2014-2-12_final508.pdf?AspxAutoDetectCookieSupport=1.
- Magasi, S., Ryan, G., Revicki, D., Lenderking, W., Hays, R. D., Brod, M., et al. (2012). Content validity of patient-reported outcome measures: Perspectives from a PROMIS meeting. *Quality of Life Research*, *21*(5), 739–746.
- Kratz, A. L., Schilling, S. G., Goesling, J., & Williams, D. A. (2015). Development and initial validation of a brief self-report measure of cognitive dysfunction in fibromyalgia. *The Journal of Pain*, *16*(6), 527–536.