



Assessing sleep and pain among adults with myalgic encephalomyelitis/chronic fatigue syndrome: psychometric evaluation of the PROMIS® sleep and pain short forms

Manshu Yang¹ · San Keller² · Jin-Mann S. Lin³

Accepted: 7 July 2022

© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

Abstract

Purpose To evaluate the psychometric properties of the patient-reported outcome measurement information system® (PROMIS) short forms for assessing sleep disturbance, sleep-related impairment, pain interference, and pain behavior, among adults with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS).

Methods Data came from the Multi-Site ME/CFS study conducted between 2012 and 2020 at seven ME/CFS specialty clinics across the USA. Baseline and follow-up data from ME/CFS and healthy control (HC) groups were used to examine ceiling/floor effects, internal consistency reliability, differential item functioning (DIF), known-groups validity, and responsiveness.

Results A total of 945 participants completed the baseline assessment (602 ME/CFS and 338 HC) and 441 ME/CFS also completed the follow-up. The baseline mean T-scores of PROMIS sleep and pain measures ranged from 57.68 to 62.40, about one standard deviation above the national norm (T-score = 50). All four measures showed high internal consistency ($\omega = 0.92$ to 0.97) and no substantial floor/ceiling effects. No DIF was detected by age or sex. Known-groups comparisons among ME/CFS groups with low, medium, and high functional impairment showed significant small-sized differences in scores ($\eta^2 = 0.01$ to 0.05) for the two sleep measures and small-to-medium-sized differences ($\eta^2 = 0.01$ to 0.15) for the two pain measures. ME/CFS participants had significantly worse scores than HC ($\eta^2 = 0.35$ to 0.45) for all four measures. Given the non-interventional nature of the study, responsiveness was evaluated as sensitivity to change over time and the pain interference measure showed an acceptable sensitivity.

Conclusion The PROMIS sleep and pain measures demonstrated satisfactory psychometric properties supporting their use in ME/CFS research and clinical practice.

Keywords Sleep · Pain · Internal consistency reliability · Differential item functioning · Known-groups validity · Responsiveness · Myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS)

Introduction

Myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) is a debilitating long-term condition that has affected approximately 836,000 to 2.5 million Americans [1–4] yet remains poorly understood by the healthcare community [5,

6]. Individuals with moderate-to-severe ME/CFS are often confined to their homes, while those with very severe disease are mostly bedbound [7, 8]. Consequently, the illness poses tremendous burdens on patients, their caregivers, and the society, costing the US economy \$18–\$51 billion annually [9–11].

ME/CFS is characterized by inability to perform usual activities and profound fatigue that lasts for 6 months or longer. In addition, post-exertional malaise, sleep problems, and either orthostatic intolerance or cognitive problems are required for diagnosis. Additional symptoms may include pain, headaches, and gastrointestinal issues [12, 13]. Recent research has found that symptoms of Long COVID-19 or post-COVID-19 conditions were similar to those of ME/CFS

✉ Manshu Yang
myang@uri.edu

¹ Department of Psychology, University of Rhode Island, 142 Flagg Road, Kingston, RI 02881, USA

² American Institutes for Research, Chapel Hill, NC, USA

³ Centers for Disease Control and Prevention, Atlanta, GA, USA

[14–18]; thus, ME/CFS research may provide insights useful to the study of Long COVID-19 and vice versa.

Identifying valid measures to characterize and track ME/CFS is a critical step toward understanding this illness and other post-infectious fatiguing conditions. The multi-site clinical assessment of myalgic encephalomyelitis/chronic fatigue syndrome (MCAM) study [6], aiming to improve how ME/CFS symptoms and their impact on quality of life could be measured, collected data using the Patient-Reported Outcome Measurement Information System® (PROMIS®) short form measures. While an earlier study supported the validity of the PROMIS Fatigue short form for ME/CFS [19], the illness presentation is known to be broader than the fatigue domain alone.

Research and clinical case definitions for ME/CFS [12, 20–23] consistently identify sleep problems as one of the required symptoms for ME/CFS diagnosis, thus it is crucial to identify and evaluate a measure that assesses both sleep quality and its impact on quality of life for ME/CFS, with standardized scores comparable across studies. Findings have been mixed regarding the type and severity of sleep problems in ME/CFS [24, 25], largely due to a lack of standard measures that can be consistently adopted. Research using various self-report measures found that 87–95% of ME/CFS patients reported unrefreshing sleep [1, 25–28], and ME/CFS patients showed significantly poorer sleep quality and more daytime dysfunction compared to healthy controls [29–31]. In contrast, studies implementing objective sleep measures [32] often did not observe more sleep-related difficulties in those with ME/CFS than in healthy controls [33–35].

Pain is another characteristic of ME/CFS listed as either a required or additional symptom in case definitions and commonly reported by patients [36, 37]. Research found that 80–94% of people with ME/CFS experience some type of pain [1, 38–40]. Importantly, other than fatigue, pain has been identified as the most troublesome ME/CFS symptom, with 65% of severely ill patients identifying pain as the top three most troublesome [14]. Among people with ME/CFS, research found that pain was associated with reduced functioning and quality of life [41]; if accompanied with depression, pain was also associated with anxiety [42]. Several measures have been used in clinical settings to evaluate ME/CFS pain [43–49]; however, such measures often focus on the frequency and severity of pain [37], while the specific consequences of pain on daily functioning and interference with life have not been fully assessed.

The primary goal of the present study is to evaluate the psychometric properties of four PROMIS® short forms to fully describe the experiences of sleep problems and pain for people suffering from ME/CFS. Using item response theory (IRT), the PROMIS short forms produce precise scores while reducing respondents' burden. Moreover, the

standardized PROMIS scores enable comparisons across studies or patient populations, which help bring meanings to the scores in both clinical and research settings.

Methods

Data source and study sample

Data came from the MCAM study [6]—a multiple-stage study with a rolling cohort design to examine the heterogeneity in patients. MCAM participants were recruited between 2012 and 2020 from ME/CFS specialty clinics across seven US states (CA, FL, NC, NJ, NV, NY, and UT). Patient eligibility was based on ME/CFS expert clinician diagnosis of the illness. Participants were aged 18–70 years at their baseline enrollment and had been diagnosed with CFS, ME, or post-infectious fatigue or managed as other ME/CFS patients in the clinical practice. The study was approved by the Institutional Review Boards of the Centers for Disease Control and Prevention and participating clinics.

Of 602 participants with ME/CFS who completed the baseline assessment, 441 also completed a follow-up assessment approximately 10 to 14 months later. No specific intervention was delivered to participants between the baseline and follow-up assessments. Most of the analyses reported here were conducted using baseline data from ME/CFS participants, whereas both baseline and follow-up data were used to evaluate the responsiveness of the measures over time. In addition, baseline data from 338 healthy controls (HC) were used to examine known-groups validity.

Measures

Four PROMIS short forms related to sleep and pain were administered: web-based platforms were used at five clinics and paper forms at the other two clinics. Based on previous literature [50], we assumed comparable responses between electronic surveys and paper-and-pencil surveys. PROMIS was developed using a mixed-method approach and calibrated all items to 2000 US Census population using IRT methods [51] on a T-score metric where a score of 50 represents the mean score of the US general population and 10 is the standard deviation [52].

PROMIS sleep short forms

We administered PROMIS v1.0 sleep measures: the *Sleep Disturbance Short Form 8b* (SD-SF) and the *Sleep-Related Impairment Short Form 8a* (SRI-SF), both with eight items [53]. All items have response options on a five-point Likert scale. Four items from SD-SF and one item from SRI-SF

were reverse coded so that higher scores indicate greater sleep disturbance or related impairment.

PROMIS pain short forms

We administered PROMIS v1.0 pain measures: the *Pain Interference Short Form 6b* (PI-SF, six items with five-point Likert scale) [54] and the *Pain Behavior Short Form 7a* (PB-SF, seven items with six-point Likert scale) [55]. For all items, higher scores indicate more impairment.

Patient characteristics

We collected data on the number of hours spent in vertical or horizontal activities per day and physical health measured by the SF-36 Health Survey (SF-36). These questions were administered to assess ME/CFS functional impairment along with the PROMIS sleep and pain measures at baseline and follow-up. We used these measures to define groups of participants differing in functional impairment for evaluating known-groups validity and responsiveness. Additionally, age and sex were used in the differential item functioning (DIF) analysis.

Analyses

Analyses were conducted using SAS Version 9.4 [56], Mplus Version 8.6 [57], and IRTPRO Version 5.1 [58].

Descriptive statistics

For each PROMIS sleep/pain short form, we calculated the mean and standard deviation of raw sum scores and examined the proportion of participants with ME/CFS at the lowest or highest possible score. A floor or ceiling effect was defined as 15% or more of respondents having the lowest or highest score [59, 60].

Unidimensionality

We evaluated the scale unidimensionality by fitting a 1-factor confirmatory factor analysis (CFA) model separately to each sleep/pain short form, assuming categorical indicators and using the WLSMV estimator [57]. If model fit was poor, exploratory factor analysis (EFA) and bi-factor CFA were further conducted to explore whether scales were at least essentially unidimensional [61–63]. Model fit was tested, with the Comparative Fit Index (CFI) > 0.95, the Tucker–Lewis Index (TLI) > 0.95, and the Root Mean Square Error of Approximation (RMSEA) < 0.06 considered as good fit [64].

IRT scoring

We used two-parameter, graded response models [65] (a type of IRT model) to obtain PROMIS T-scores for participants with ME/CFS and HC at baseline and at follow-up. Item parameters were fixed at the published values from PROMIS item banks v1.0, which were calibrated using a large sample representing 2000 US Census population. We employed response pattern scoring using IRTPRO Version 5.1 [58], based on item parameters available on Assessment Center (<https://www.assessmentcenter.net/>).

Reliability of scores

Internal consistency reliability for each of the four PROMIS sleep and pain measures was evaluated using Cronbach's standardized alpha coefficient, categorical omega coefficient [66], and the amount of measurement error in the T-score under IRT. We computed omega based on item loadings and thresholds from 1-factor or bi-factor CFA models via Bayesian estimation [67]. The omega coefficient provides a more accurate estimate of internal consistency than alpha because it makes more realistic assumptions about the measurement models (e.g., does not assume equal factor loadings of all items or uncorrelated error variances) [68–70]. Although a reliability coefficient > 0.70 (or a T-score standard error of measurement [T-SEM] < 5.5) is considered acceptable for group-level analyses, a reliability coefficient > 0.90 (or a T-SEM < 3.2) is needed for precisely assessing individual patients [71, 72]. We hypothesized that, consistent with previous findings with other health conditions [73–77], *the PROMIS sleep and pain scores among participants with ME/CFS would have reliability estimates exceeding those recommended for individual-level comparisons (i.e., omega > 0.90 and T-SEM < 3.2).*

DIF analysis

We conducted DIF analysis to detect potential measurement bias across population subgroups differing in sex and age. Evidence of DIF occurs when respondent subgroups (e.g., male vs. female) differ in their probabilities of endorsing an item response category after controlling for the underlying trait being measured. DIF suggests that item score differences between subgroups (e.g., male vs. female) may be merely due to group membership or different interpretations of the item content, rather than reflecting true differences in the trait being measured.

We examined the possibility of DIF by sex and age for each item in the four PROMIS sleep and pain short forms using the Wald test [78, 79]. The measurement properties of each item were compared across three age groups: 18–39, 40–59, and 60 or above, which allowed a sufficiently large

number of participants in each group to support this analysis. Patients ages 40–59 were specifically combined as a category since ME/CFS is more prevalent in this age range [80, 81]. Two DIF comparisons were made for ages 18–39 vs. ages 40+ and ages 40–59 vs. ages 60+. For the Wald test, a non-significant χ^2 value indicates no detectable DIF. We used the Benjamin–Hochberg [82, 83] procedure to control for the multiple comparisons involved in checking DIF for each item by sex and age. *We hypothesized that there would be no evidence of DIF for items from the four PROMIS sleep and pain short forms in this sample of participants with ME/CFS.*

Known-groups validity

Known-groups validity of each PROMIS sleep/pain short form was evaluated by comparing T-scores for groups that are known to differ in their trait levels. *We hypothesized that the ME/CFS group with higher levels of functional impairment would have PROMIS scores indicating greater sleep disturbance/impairment or pain interference/behavior; and that the ME/CFS group would exhibit greater sleep disturbance/impairment or pain interference/behavior than HC.* We measured the level of functional impairment based on three variables, including (1) hours spent in *vertical* activities (e.g., sitting, standing, or walking) per day, (2) hours spent in *horizontal* activities (e.g., resting in recliner with feet up, napping, sleeping in bed) per day, and (3) overall *physical health* (measured by the Physical Component Summary [PCS] T-scores of the SF-36). Using three different types of measures allows us to better define the known-groups with the former two measures specific to ME/CFS and commonly used by ME/CFS expert clinicians [84], while the third measure capturing generic functional status. Fewer hours of vertical activities, more hours of horizontal activities, and lower SF-36 PCS scores indicate more functional impairment. Additional detail regarding these measures is provided in the supplement materials of [6] and [19].

For vertical activities and physical health, participants with ME/CFS were divided into three similar-sized groups with “low,” “medium,” and “high” level of impairment based on tertiles. For horizontal activities, we divided ME/CFS participants into only two groups (15 h vs. < 15 h of horizontal activities per day), because too many participants (about 47%) were at the “ceiling” of 15 h of horizontal activities per day [19].

Analysis of variance (ANOVA) was used to examine mean differences in PROMIS T-scores for SD-SF, SRI-SF, PI-SF, and PB-SF among known-groups defined by the three aforementioned variables for participants with ME/CFS and between the ME/CFS and HC groups. The Tukey–Kramer method [85] was adopted for multiple comparison adjustment among known-groups. Known-groups

validity was considered acceptable when the difference in mean T-scores across groups was significant at $\alpha=0.05$. We interpreted the size of these differences using η^2 (dividing the sum of squares for the known-groups effect by the total sum of squares): following convention, η^2 around 0.01, 0.09, and 0.25 represents small, medium, and large effect, respectively [86, 87].

Responsiveness

Responsiveness, or sensitivity to change, represents how well a measure’s scores reflect changes over time when true changes occur. *We hypothesized that participants with ME/CFS with “improved,” “stable,” and “worsened” health would show a significant decrease, no significant changes, and a significant increase in their PROMIS sleep/pain scores, respectively.* We initially categorized ME/CFS participants into three groups of “improved,” “stable,” and “worsened” using the three aforementioned measures of functional status. As detailed in [19], for vertical/horizontal activities, “stable” participants were defined as those who had < 1 h of change from baseline to follow-up, because time spent in these activities was reported in integer hours and there was no established threshold for the minimal clinically important difference (MCID). In terms of SF-36 PCS T-score, “stable” participants were those with ≤ 5 points of change. Previous literature suggests MCIDs of 2.5 to about 7 for SF-36 PCS across different patient populations [88–90], but its MCID for ME/CFS has not been established. Therefore, we used the half standard deviation approach [91] and chose an MCID of 5 as the threshold of categorizing participants when assessing the responsiveness.

For horizontal activities, we further combined the “stable” and “worsened” groups into a group of “not improved” and compared those who “improved” versus “not improved.” This is because over 77% of “stable” participants were at the worst possible functional status (i.e., 15 h of daily horizontal activities) at both baseline and follow-up. Such participants may not be truly stable, as they might have experienced a worsening in horizontal activities that was undetectable.

We used ANOVA to examine if changes in PROMIS sleep/pain scores significantly differed among the “improved,” “stable,” and “worsened” groups (or between the “improved” and “not improved” groups for horizontal activities). Additionally, we calculated Guyatt’s responsiveness statistic (GRS) to describe the effect size comparing the “improved” groups to the “stable,” “worsened,” or “not improved” groups. The GRS is the mean change in PROMIS sleep/pain scores for the target group (i.e., “improved”) divided by the standard deviation of the comparison group (e.g., “not improved”) [92] and is interpreted as small (≥ 0.2 and < 0.5), medium (≥ 0.5 and < 0.8), and large (≥ 0.8) [86].

Results

Most participants with ME/CFS were female (72.6%), white (88.4%), and not currently working (70.1%) (Table 1). Their mean age was 48.0 years, with an average illness duration of 14.2 years and more than half of them had sudden illness onset.

Descriptive statistics

Table 2 summarizes the distribution of raw sum scores for each PROMIS sleep/pain short form among participants

with ME/CFS. The proportions of participants at the highest possible or lowest possible impairment score were all below the threshold of 15%, suggesting no substantial floor/ceiling effects. Approximately 10% of participants reported “not at all/never” on all the pain interference items or reported “I had no pain” on all the pain behavior items, which indicated that a subgroup of ME/CFS participants may not have suffered from pain at the time that they answered the questions.

Unidimensionality

As shown in Table 3, the CFI and TLI demonstrated excellent fit (> 0.95) of the pain short forms to a unidimensional

Table 1 Sample characteristics at baseline

	ME/CFS (<i>n</i> = 602)		HC (<i>n</i> = 338)	
	<i>n</i>	%	<i>n</i>	%
Sex				
Female	437	72.6	222	65.7
Male	165	27.4	115	34.0
Missing	0	0	1	0.3
Race				
White	532	88.4	192	56.8
Black/African American	12	2.0	22	6.5
All others	28	4.6	88	26.0
Missing	30	5.0	36	10.7
Marital status				
Married/committed	316	52.5	168	49.7
Previously married	104	17.3	60	17.7
Never married	171	28.4	101	29.9
Missing	11	1.8	9	2.7
Employment				
Full-time	91	15.1	177	52.4
Part-time	63	10.5	60	17.8
Not working	422	70.1	89	26.3
Missing	26	4.3	12	3.5
Educational attainment				
Less than high school	4	0.7	5	1.5
High school graduate	131	21.8	98	29.0
College graduate	237	39.4	126	37.3
Post college	211	35.0	100	29.6
Missing	19	3.1	9	2.6
Illness onset status				
Gradual	192	31.9	N.A	N.A
Sudden	330	54.8	N.A	N.A
Missing	80	13.3	N.A	N.A
	Mean	SD	Mean	SD
Age, years	48.0	13.0	43.3	15.2
Illness Duration, years	14.2	10.0	N.A	N.A

ME/CFS myalgic encephalomyelitis/chronic fatigue syndrome, HC healthy control, SD standard deviation

Table 2 Measure-level raw sum score distributions for PROMIS short forms of sleep disturbance, sleep-related impairment, pain interference, and pain behavior, among participants with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) at baseline from the multi-site clinical assessment of ME/CFS (MCAM) study

Measure	N	Mean	SD	Participants at the lowest possible raw sum score		Participants at the highest possible raw sum score	
				n	%	n	%
PROMIS sleep disturbance (raw sum score 8–40)	583	28.30	6.94	1	0.2	24	4.0
PROMIS sleep-related impairment (raw sum score 8–40)	577	26.67	7.13	2	0.3	21	3.5
PROMIS pain interference (raw sum score 6–30)	585	18.83	7.49	61	10.1	35	5.8
PROMIS pain behavior (raw sum score 7–42)	578	23.68	7.46	60	10.0	0	0.0

SD standard deviation

Table 3 Model fit indices from the confirmatory factor analyses (CFA) for PROMIS short forms of sleep disturbance, sleep-related impairment, pain interference, and pain behavior, among participants with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) at baseline from the multi-site clinical assessment of ME/CFS (MCAM) study

Measure	CFA model	Model Fit Indices		
		CFI	TLI	RMSEA
PROMIS sleep disturbance	1-Factor	0.922	0.890	0.231
	Bi-factor	0.999	0.997	0.040
PROMIS sleep-related impairment	1-Factor	0.942	0.919	0.247
	Bi-factor	0.999	0.998	0.042
PROMIS pain interference	1-Factor	0.997	0.995	0.173
PROMIS pain behavior	1-Factor	0.992	0.988	0.113
<i>Threshold of good fit</i>		<i>> 0.95</i>	<i>> 0.95</i>	<i>< 0.06</i>

CFI Comparative Fit Index; TLI Tucker–Lewis Index; RMSEA root mean square error of approximation

The values in italics are “thresholds” used to evaluate the model fit indices values reported

CFA model but marginal fit (> 0.90) of the sleep short forms. Therefore, we further conducted EFA and bi-factor analyses for the two sleep short forms and found that all model fit indices were greatly improved to the acceptable ranges. Although the RMSEAs for the pain measures were beyond the published criterion (> 0.06), we considered them as sufficiently unidimensional based on high CFI and TLI values. Moreover, the RMSEA estimates from the current study were consistent with those previously reported for PROMIS measures [54, 55, 93].

IRT scoring

The mean IRT-based T-scores of the four measures ranged from 57.68 to 62.40 for participants with ME/CFS at baseline (see Table 4). As such, the four measures had mean scores about one standard deviation (10 points on the T-score metric) above the national norm (T-score = 50). These scores indicate greater pain and sleep problems compared to the general population, consistent with clinical understanding of

the illness and, thus, supporting the validity of using them for ME/CFS.

Reliability

As shown in Table 4, among participants with ME/CFS, all four PROMIS sleep/pain measures showed high internal consistency with ω ranging from 0.92 to 0.97. Under the IRT framework, the average standard errors of T-scores ranged from 2.1 to 2.7 (see Table 4), which were well below the hypothesized threshold of 3.2 (corresponding to a reliability of 0.9). Compared to Cronbach’s α (ranging from 0.88 to 0.97 in this study), the ω coefficient and the average standard error of IRT T-score are computed under a more lenient and realistic assumption, allowing each item to be linked to the underlying construct (e.g., pain interference) to differing degree. Therefore, we considered the ω coefficient or the standard error of T-score a more accurate estimate of reliability. Although the α coefficients were slightly below the hypothesized threshold of 0.9 for the two sleep measures, these estimates may reflect an underestimation of internal consistency, as the tau-equivalence assumption (i.e., equal

Table 4 Measure-Level T-score distributions and reliability estimates for PROMIS short forms of sleep disturbance, sleep-related impairment, pain interference, and pain behavior, among participants with

myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) at baseline from the multi-site clinical assessment of ME/CFS (MCAM) study

Measure	<i>n</i>	IRT T-score		Reliability		
		Mean	SD	Cronbach's α	ω^a	SEM
PROMIS sleep disturbance	602	59.17	8.35	0.88	0.92	2.7
PROMIS sleep-related impairment	601	62.40	8.73	0.89	0.93	2.5
PROMIS pain interference	601	62.02	9.84	0.97	0.97	2.3
PROMIS pain behavior	601	57.68	8.93	0.92	0.92	2.1

IRT item response theory, SD standard deviation, SEM standard error of measurement, based on IRT T-scores

^aCategorical ω were computed using Green and Yang's approach [66], based on item loadings and thresholds estimates from 1-factor CFA models (i.e., total omega) for the pain interference and pain behavior short forms and from bi-factor CFA models (i.e., hierarchical omega) for the sleep disturbance and sleep-related impairment short forms

factor loadings for all items) required for computing α was violated, with loadings ranging from 0.62 to 0.89 for sleep disturbance and from 0.57 to 0.93 for sleep-related impairment. Taken together, the four PROMIS measures provided highly reliable scores not only for group-level analysis but also for assessing individual ME/CFS patients.

DIF

Table 5 shows the Wald test results for detecting potential DIF by sex and by age. Before using the Benjamini–Hochberg correction for multiplicity, two PROMIS SD-SF items (“I was satisfied with my sleep” and “My sleep was refreshing”) and one item from PROMIS PB-SF (“When I was in pain I grimaced”) showed possible DIF by age. Another item from PROMIS PI-SF (“How much did pain interfere with your enjoyment of recreational activities?”) showed possible DIF by sex, with a p value < 0.05 . However, after correction for multiplicity, none of the items exhibited significant DIF by sex or by age.

Known-groups validity

Results in Tables 6 and 7 show that the omnibus hypothesis of no differences among known-groups was rejected with $p < 0.05$ for each of the four PROMIS measures (see table footnotes for details), providing supportive evidence for the validity of the sleep and pain short forms for ME/CFS. However, for the two sleep-related PROMIS measures, although mean scores appeared to increase monotonically across low, medium, and high functional impairment groups, the differences between medium and high functional impairment groups were not statistically significant (see Table 6 and its footnotes). For the two pain-related PROMIS measures, we found significant differences for all pairwise comparisons between groups defined by SF-36 PCS scores, with a monotonic increase in mean scores across low, medium,

and high functional impairment groups (see Table 7 and its footnotes). When groups were defined by vertical activities, however, we only found significant difference between low and high functional impairment groups but could not differentiate the medium impairment group from the other two groups (see Table 7 and its footnotes). Participants with ME/CFS had significantly higher (i.e., worse) mean scores than HC for all four PROMIS measures. For the two sleep-related PROMIS measures, effect sizes, η^2 , were small for known-groups defined by functional impairment (i.e., vertical/horizontal activities and SF-36 PCS scores). For the two pain-related PROMIS measures, small-sized and medium-sized effects were found for known-groups defined by vertical/horizontal activities and SF-36 PCS scores, respectively. When comparing ME/CFS to HC participants, effect sizes were large for all four PROMIS measures.

Responsiveness

For the two sleep-related PROMIS short forms, no significant differences were found among groups of participants with ME/CFS defined by whether they improved, remained stable, or worsened based on SF-36 PCS scores or horizontal activity. Overall, sleep-related impairment change scores significantly differed across groups defined by vertical activity ($F = 6.7$, $p = 0.001$), but in pairwise group comparisons, the “stable” group did not significantly differ from either the “improved” or “worsened” group. Therefore, there was insufficient evidence to support the responsiveness of the PROMIS sleep measures for ME/CFS.

For the two pain-related PROMIS short forms, significant overall differences were found among groups of participants with ME/CFS defined by whether they improved, remained stable, or worsened with respect to SF-36 PCS scores. In pairwise group comparisons defined by SF-36 PCS scores, we found that all three groups significantly differed from each other for pain interference; whereas,

Table 5 Differential item functioning (DIF) statistics by sex and by age for PROMIS short forms of sleep disturbance, sleep-related impairment, pain interference, and pain behavior, based on baseline myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) participant data (n = 602) from the multi-site clinical assessment of ME/CFS (MCAM) study

Label	DIF by Sex Male vs Female			DIF by Age 18–39 vs. 40+			DIF by Age 40–59 vs. 60+		
	χ^2	df	p	χ^2	df	p	χ^2	df	p
PROMIS sleep disturbance short form									
Item1: restless	1.1	5	0.957	4	5	0.549	11	5	0.051
Item2: satisfied sleep	4.3	5	0.507	4.9	5	0.425	11.4	5	0.044
Item3: refreshing	0.9	5	0.973	2.5	4	0.639	16.5	4	0.002
Item4: falling asleep	0.9	5	0.971	2.8	5	0.726	7.7	5	0.176
Item5: staying asleep	5.9	5	0.314	7.5	5	0.185	5.9	5	0.313
Item6: trouble sleeping	7.2	5	0.204	4	5	0.550	7.8	5	0.169
Item7: enough sleep	8.8	5	0.116	3.5	5	0.620	7.9	5	0.163
Item8: sleep quality	3	5	0.702	2.4	4	0.662	6.1	4	0.192
PROMIS sleep-related impairment short form									
Item1: getting things done	4.9	5	0.429	1.5	5	0.918	5	5	0.423
Item2: alert when woke up	9.4	5	0.093	1.4	5	0.919	4.5	5	0.485
Item3: tired	5.5	5	0.355	1.4	4	0.843	7.7	4	0.104
Item4: problems during day	1.5	5	0.911	3.2	5	0.667	2	5	0.845
Item5: hard time concentrating	4.9	5	0.430	7.4	5	0.189	1.3	5	0.938
Item6: irritable	2.3	5	0.813	3.4	5	0.634	3.5	5	0.621
Item7: sleepy during daytime	0.4	5	0.997	4	5	0.555	9.9	5	0.079
Item8: trouble staying awake	8.2	5	0.144	7.4	5	0.192	3.8	5	0.581
PROMIS pain interference short form									
Item1: enjoyment of life	5.2	5	0.398	10	5	0.076	6.4	5	0.271
Item2: ability to concentrate	4.1	5	0.539	7.1	5	0.215	6.2	5	0.290
Item3: day-to-day activities	3.6	5	0.604	1.4	5	0.924	4.6	5	0.461
Item4: recreational activities	15.8	5	0.008	1.6	5	0.898	1.7	5	0.888
Item5: tasks away from home	6.2	5	0.286	5.3	5	0.380	5.5	5	0.358
Item6: socializing with others	5	5	0.421	1.9	5	0.861	7.1	5	0.211
PROMIS pain behavior short form									
Item1: irritable	2.4	6	0.874	10.9	6	0.092	6.9	6	0.330
Item2: grimaced	3.9	6	0.689	2.6	6	0.863	14.8	6	0.022
Item3: moved extremely slowly	10.7	6	0.096	9.3	6	0.159	7.9	6	0.249
Item4: moved stiffly	3.5	6	0.749	6.3	6	0.388	3.1	6	0.793
Item5: called out for help	5.6	6	0.470	8.9	5	0.113	8.5	5	0.130
Item6: isolated from others	4.6	6	0.597	3.8	6	0.710	3.8	6	0.706
Item7: thrashed	0.6	6	0.996	7.9	5	0.163	3.5	5	0.619

only the difference between “improved” and “worsened” groups was significant for pain behavior. It is worth noting that although significant difference was found among groups defined by vertical activities for pain behavior, the difference was not in the expected direction and indicated greater decrease in pain behavior in the “stable” group than in the “improved” group. Therefore, the significant difference for pain behavior across vertical activity groups should not be considered as evidence supporting its responsiveness.

Guyatt’s responsiveness statistics were generally small, except for pain interference that showed medium-to-large differences in change scores across “improved,” “stable,”

and “worsened” groups defined by SF-36 PCS scores. To summarize, only the pain interference short form showed strong evidence supporting its responsiveness among ME/CFS participants (Tables 8 and 9).

Discussion

The PROMIS sleep and pain short forms are generic (i.e., not condition-specific) measures that have been tested in various patient populations (e.g., fibromyalgia, hypertension, sleep disorder, rheumatoid arthritis, sickle cell disease) [53, 77, 94, 95]. The current study demonstrated that

Table 6 Mean T-scores of the PROMIS sleep disturbance (SD) and sleep-related impairment (SRI) short forms (SF), by three indicators of functional impairment level as well as between participants with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) and healthy controls (HC) from the multi-site clinical assessment of ME/CFS (MCAM) study

Known-groups	PROMIS SD-SF ^a				PROMIS SRI-SF ^b					
	<i>n</i>	Mean	SD	<i>F</i>	η^2 ^c	<i>n</i>	Mean	SD	<i>F</i>	η^2 ^c
ME/CFS functional impairment level defined by hours of vertical activities per day										
Low (≥ 10 h)	212	58.1	8.9	4.0	0.014	212	60.6	9.2	7.9	0.026
Medium (≥ 5 and < 10 h)	187	58.9	7.6			187	62.7	8.4		
High (< 5 h)	184	60.4	8.1			184	64.0	8.1		
ME/CFS functional impairment level defined by hours of horizontal activities per day										
Lower (< 15 h)	254	58.3	8.6	4.1	0.007	254	61.0	9.3	11.6	0.020
Higher (15 h)	330	59.7	8.0			330	63.4	8.1		
ME/CFS functional impairment level defined by SF-36 PCS score										
Low (T-score ≥ 28.78)	196	57.2	8.6	9.2	0.030	196	59.8	9.3	13.8	0.045
Medium ($20.49 \leq$ T-score < 28.78)	202	59.5	7.8			202	63.7	8.1		
High (T-score < 20.49)	195	60.8	8.3			194	63.8	8.2		
ME/CFS vs. HC										
ME/CFS	602	59.2	8.3	512.1	0.353	601	62.4	8.7	733.3	0.439
HC	338	45.7	9.5			338	44.6	11.2		

SD standard deviation

^aOverall differences across groups were significant at $p < 0.05$, with groups defined by vertical/horizontal activities, SF-36 PCS scores, and ME/CFS vs. healthy controls (HC). For pairwise comparisons between any two functional impairment levels of vertical activities or SF-36 physical component summary (PCS) scores: only difference of low vs. high functional impairment, defined by vertical activities, was significant ($p < 0.05$); only differences of low vs. medium and low vs. high functional impairment, defined by SF-36 PCS scores, were significant ($p < 0.05$)

^bOverall differences across groups were significant at $p < 0.001$, with groups defined by vertical/horizontal activities, SF-36 PCS scores, and ME/CFS vs. healthy controls (HC). For pairwise comparisons between any two functional impairment levels of vertical activities or SF-36 PCS scores, only differences of low vs. medium and low vs. high functional impairment were significant ($p < 0.05$)

^c η^2 is an effect size measure and was computed by dividing the sum of squares for the known-groups effect by the total sum of squares

Table 7 Mean T-scores of the PROMIS pain interference (PI) and pain behavior (PB) short forms (SF), by three indicators of functional impairment level as well as between participants with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) and healthy controls (HC) from the multi-site clinical assessment of ME/CFS (MCAM) study

Known-groups	PROMIS PI-SF ^a				PROMIS PB-SF ^b					
	<i>n</i>	Mean	SD	<i>F</i>	η^2 ^c	<i>n</i>	Mean	SD	<i>F</i>	η^2 ^c
ME/CFS functional impairment level defined by hours of vertical activities per day										
Low (≥ 10 h)	212	60.0	10.2	7.4	0.025	212	56.1	9.6	5.0	0.017
Medium (≥ 5 and < 10 h)	187	62.0	9.3			187	58.0	8.4		
High (< 5 h)	184	63.8	9.7			184	58.8	8.6		
ME/CFS functional impairment level defined by hours of horizontal activities per day										
Lower (< 15 h)	254	60.9	10.4	4.5	0.008	254	56.6	9.8	5.3	0.009
Higher (15 h)	330	62.6	9.4			330	58.3	8.2		
ME/CFS functional impairment level defined by SF-36 PCS score										
Low (T-score ≥ 28.78)	196	57.0	10.4	50.8	0.147	196	53.5	10.8	40.2	0.120
Medium ($20.49 \leq$ T-score < 28.78)	202	62.8	9.2			202	58.5	8.3		
High (T-score < 20.49)	194	66.2	7.6			194	61.0	4.9		
ME/CFS vs. HC										
ME/CFS	601	62.0	9.8	749.8	0.445	601	57.7	8.9	630.3	0.403
HC	335	45.2	7.2			335	41.5	10.4		

SD standard deviation

^aOverall differences across groups were significant at $p < 0.05$, with groups defined by vertical/horizontal activities, SF-36 PCS scores, and ME/CFS vs. healthy controls (HC). For pairwise comparisons between any two functional impairment levels of vertical activities or SF-36 PCS scores: only difference of low vs. high functional impairment, defined by vertical activities, was significant ($p < 0.001$); differences of low vs. medium, low vs. high, and medium vs. high functional impairment, defined by SF-36 PCS scores, were all significant ($p < 0.001$)

^bOverall differences across groups were significant at $p < 0.05$, with groups defined by vertical/horizontal activities, SF-36 PCS scores, and ME/CFS vs. HC. For pairwise comparisons between any two functional impairment levels of vertical activities or SF-36 PCS scores: only difference of low vs. high functional impairment, defined by vertical activities, was significant ($p < 0.01$); differences of low vs. medium, low vs. high, and medium vs. high functional impairment, defined by SF-36 PCS scores, were all significant ($p < 0.01$)

^c η^2 is an effect size measure and was computed by dividing the sum of squares for the known-groups effect by the total sum of squares

Table 8 Mean changes in T-scores based on PROMIS sleep disturbance (SD) and sleep-related impairment (SRI) short forms (SF), among “improved,” “stable,” and “worsened” groups (or between “improved” and “not improved” groups) defined by three anchor measures from the multi-site clinical assessment of ME/CFS (MCAM) study

Change status	PROMIS SD-SF						PROMIS SRI-SF					
	<i>n</i>	Mean	SD	<i>F</i> ^a	Guyatt's Responsiveness Statistic		<i>n</i>	Mean	SD	<i>F</i> ^a	Guyatt's Responsiveness Statistic	
					Improved vs Worsened /Not Improved	Stable vs Worsened					Improved vs Worsened /Not Improved	Stable vs Worsened
Change in hours of vertical activities per day, from baseline to follow-up												
Improved	150	-1.07	7.19	2.7	-0.253	-0.030	149	-1.41	7.97	6.7	-0.425	-0.220
Stable	103	0.45	6.83				103	0.02	6.29			
Worsened	169	0.66	6.82				169	1.55	6.96			
Change in hours of horizontal activities per day, from baseline to follow-up												
Improved	114	-1.07	7.02	3.7	-0.212		114	-0.34	8.18	0.7	-0.095	
Not improved	308	0.41	6.97				307	0.32	6.93			
Change in SF-36 physical component summary (PCS) score, from baseline to follow-up												
Improved	107	-1.24	6.50	2.9	-0.362	-0.142	107	-0.71	7.94	1.1	-0.230	-0.101
Stable	261	0.29	7.02				260	0.22	6.80			
Worsened	55	1.28	6.94				55	0.94	7.13			

SD standard deviation

^aBold values indicate significant overall differences across groups at *p* < 0.05

Table 9 Mean changes in T-scores based on PROMIS pain interference (PI) and pain behavior (PB) short forms (SF), among “improved,” “stable,” and “worsened” groups (or between “improved” and “not improved” groups) defined by three anchor measures from the multi-site clinical assessment of ME/CFS (MCAM) study

Change status	PROMIS PI-SF				PROMIS PB-SF							
	<i>n</i>	Mean	SD	<i>F</i> ^a	<i>n</i>	Mean	SD	<i>F</i> ^a				
	Guyatt's Responsiveness Statistic				Guyatt's Responsiveness Statistic							
				Improved vs Worsened /Not Improved	Improved vs Worsened /Not Improved			Improved vs Worsened /Not Improved				
Change in hours of vertical activities per day, from baseline to follow-up												
Improved	150	-1.78	7.63	1.9	-0.218	-0.045	150	-0.64	6.01	3.3	-0.174	-0.318
Stable	103	-0.60	6.58				103	-1.54	7.04			
Worsened	169	-0.29	6.83				169	0.45	6.28			
Change in hours of horizontal activities per day, from baseline to follow-up												
Improved	114	-1.78	7.79	2.5	-0.182		114	-1.03	7.17	1.5	-0.142	
Not Improved	308	-0.54	6.80				308	-0.17	6.12			
Change in SF-36 PCS score, from baseline to follow-up												
Improved	107	-3.26	7.43	13.5	-0.874	-0.508	107	-1.75	6.08	4.6	-0.522	-0.279
Stable	261	-0.79	6.67				261	-0.30	6.51			
Worsened	55	2.63	6.73				55	1.37	5.98			

SD standard deviation

^aBold values indicate significant overall differences across groups at $p < 0.05$

the four PROMIS short forms had strong reliability and validity to assess sleep and pain outcomes for individuals with ME/CFS, thus they are useful tools for researchers and clinicians to examine individuals with varying levels of functional impairment due to ME/CFS as well as comparing them to those with other illnesses.

All four measures demonstrated essential unidimensionality and excellent internal consistency not only for group-level analyses but also for monitoring individuals with ME/CFS. Minimal floor/ceiling effects were shown at the total score level, suggesting that the PROMIS short forms exhibited minimal restriction on the measurement range for sleep and pain problems within the ME/CFS population. We observed that about 10% of participants had the lowest possible scores for the two pain short forms, indicating no pain or no pain interference. This is consistent with our expectation, given that pain is a common and troublesome symptom but not a required symptom for ME/CFS diagnosis. All four measures showed acceptable known-groups validity, with small-to-medium effect sizes when comparing across functional impairment groups and with large effect sizes when comparing participants with ME/CFS to HC. As expected, the T-scores generally had a monotonic increase across ME/CFS groups with low, medium, and high functional impairment, and participants with ME/CFS had significantly higher (i.e., worse) scores than HC. No DIF was detected by age or sex for any items, suggesting that all four measures provide unbiased measurement across these population subgroups.

The two sleep short forms (sleep-related impairment and sleep disturbance) and the pain behavior short form did not show sufficient evidence to support their responsiveness to change, while the pain interference short form showed good responsiveness with medium-to-large effect sizes. Previous research has estimated the minimal clinically important difference to be 2 to 6 points for the PROMIS sleep and pain measures across different patient populations [96–99]. However, given the non-interventional nature of the present study, except for pain interference, the average changes in PROMIS T-scores from baseline to follow-up were mostly below 2 points, which dampens our ability to fully evaluate responsiveness. Furthermore, given that the PROMIS Fatigue short form was found to be responsive within the same study sample of individuals with ME/CFS [19], it is possible that the improvement or worsening of functioning (as defined by vertical/horizontal activity or SF-36 PCS score) within a 10 to 14-month period were mainly driven by changes in fatigue levels, while changes due to other symptoms, such as sleep problems and pain behavior, may be less tightly linked to functioning or may take longer to affect functioning. Therefore, further research examining a longer (e.g., 2 years) follow-up period is warranted to test the responsiveness of these scores. Moreover, a better test of responsiveness of scores would be to examine changes in

scores among patients after intervention with a therapy of known efficacy.

The present study is not without limitations. First, all MCAM study participants were receiving tertiary care and may not fully represent the broader U.S. ME/CFS population. Although the large number of participants recruited from clinics across seven states should form a diverse sample for psychometric analyses, future studies that include non-tertiary care participants are needed to evaluate the stability of parameter estimates. Second, the analyses for evaluating known-groups validity and responsiveness were compromised by the measurement of functional status, defined by hours of vertical/horizontal activity per day and the SF-36 PCS scores. While these measures were highly correlated with the leading ME/CFS symptom of fatigue, their associations with sleep problems or pain may be weaker. Studies using more relevant external criterion measures (e.g., global rating of sleep quality or pain) should be conducted to further examine the validity of PROMIS short forms. In addition, a large proportion of individuals with ME/CFS began the study with the worst possible functional impairment defined by horizontal activity. Consequently, those who were truly stable could not be distinguished from those who experienced an undetectable worsening in functional status when measured by horizontal activity. More research is needed to validate the vertical/horizontal activity measures and establish their cutoffs for MCID, so that stable participants can be consistently and precisely defined. These ME/CFS-specific measures, along with other indicators with greater individual differences and more variability, would help researchers triangulate findings to better evaluate the responsiveness and the test–retest reliability.

Despite these limitations, the results of this study provide useful information about the reliability and validity of the four PROMIS sleep and pain short forms, both in general and when used in individuals with ME/CFS, in particular. This information will facilitate the selection of patient-reported outcome measures for ME/CFS and other similar illnesses (e.g., Long COVID-19) moving forward.

Conclusion

In summary, study findings support the reliability and validity of four PROMIS short forms for assessing sleep-related impairment, sleep disturbance, pain interference, and pain behavior, among individuals with ME/CFS. Such measures could be used in research and clinical settings to facilitate understanding of the symptomatology and clinical course of ME/CFS, which is an important step for evaluating treatment effect.

Acknowledgements We would like to thank the Multi-Site Clinical Assessment of ME/CFS (MCAM) study group who made the data sets available through Research Collaboration Agreement. The group included Division of High-Consequence, Pathogens and Pathology, Centers for Disease Control and Prevention, Atlanta, Georgia: Elizabeth R. Unger, Britany Helton, Yang Chen, and Monica Cornelius; Open Medicine Institute Consortium, Mountain View, CA: Andreas Kogelnik, Catt Phan, Joan Danver, Lucinda Bateman, Jennifer Bland, Charles Lapp, Wendy Springs, Richard Podell, Trisha Fitzpatrick, Daniel Peterson, and Marco Maynard; Institute for Neuro Immune Medicine, Miami, Florida: Nancy Klimas, Elizabeth Balbin, Precious Leaks-Gutierrez, and Shuntae Parnell; and Mount Sinai Beth Israel, New York, New York: Benjamin Natelson and Diana Vu.

Author contributions All authors contributed to the study conception, design, and results interpretation. Psychometric analysis was performed by MY. The first draft of the manuscript was written by MY and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Manshu Yang (MY) and San Keller (SK) were supported by the Centers for Disease Control and Prevention (CDC) contract (75D30119P06165). Jin-Mann S. Lin (JML) was supported by the CDC.

Declarations

Competing interests The authors declare no conflict of interest.

Disclaimers The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention. The authors declare no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

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

References

- Jason, L. A., Richman, J. A., Rademaker, A. W., Jordan, K. M., Plioplys, A. V., Taylor, R. R., McCreedy, W., Huang, C. F., & Plioplys, S. (1999). A community-based study of chronic fatigue syndrome. *Archives of Internal Medicine*, *159*(18), 2129–2137.
- Reeves, W. C., Jones, J. F., Maloney, E., Heim, C., Hoaglin, D. C., Boneva, R. S., Morrissey, M., & Devlin, R. (2007). Prevalence of chronic fatigue syndrome in metropolitan, urban, and rural Georgia. *Population Health Metrics*, *5*, 5. <https://doi.org/10.1186/1478-7954-5-5>
- Rowe, P. C., Underhill, R. A., Friedman, K. J., Gurwitt, A., Medow, M. S., Schwartz, M. S., Speight, N., Stewart, J. M., Vallings, R., & Rowe, K. S. (2017). Myalgic encephalomyelitis/chronic fatigue syndrome diagnosis and management in young people: A primer. *Frontiers in Pediatrics*, *5*, 121. <https://doi.org/10.3389/fped.2017.00121>
- Cortes Rivera, M., Mastronardi, C., Silva-Aldana, C. T., Arcos-Burgos, M., & Lidbury, B. A. (2019). Myalgic encephalomyelitis/chronic fatigue syndrome: A comprehensive review. *Diagnostics (Basel)*, *9*(3), 91. <https://doi.org/10.3390/diagnostics9030091>
- Lin, J.-M.S., Brimmer, D. J., Boneva, R. S., Jones, J. F., & Reeves, W. C. (2009). Barriers to healthcare utilization in fatiguing illness: A population-based study in Georgia. *BMC Health Services Research*, *9*(1), 13. <https://doi.org/10.1186/1472-6963-9-13>
- Unger, E. R., Lin, J.-M. S., Tian, H., Natelson, B. H., Lange, G., Vu, D., Blate, M., Klimas, N. G., Balbin, E. G., Bateman, L., Allen, A., Lapp, C. W., Springs, W., Kogelnik, A. M., Phan, C. C., Danver, J., Podell, R. N., Fitzpatrick, T., Peterson, D. L., ... Rajeevan, M. S. (2017). Multi-site clinical assessment of myalgic encephalomyelitis/chronic fatigue syndrome (MCAM): Design and implementation of a prospective/retrospective rolling cohort study. *American Journal of Epidemiology*, *185*(8), 617–626. <https://doi.org/10.1093/aje/kwx029>
- Pendergrast, T., Brown, A., Sunnquist, M., Jantke, R., Newton, J. L., Strand, E. B., & Jason, L. A. (2016). Housebound versus non-housebound patients with myalgic encephalomyelitis and chronic fatigue syndrome. *Chronic Illness*, *12*(4), 292–307. <https://doi.org/10.1177/1742395316644770>
- Conroy, K., Bhatia, S., Islam, M., & Jason, L. A. (2021). Homebound versus bedridden status among those with myalgic encephalomyelitis/chronic fatigue syndrome. *Healthcare (Basel)*, *9*(2), 106. <https://doi.org/10.3390/healthcare9020106>
- Reynolds, K. J., Vernon, S. D., Bouchery, E., & Reeves, W. C. (2004). The economic impact of chronic fatigue syndrome. *Cost Effectiveness and Resource Allocation*, *2*(1), 4. <https://doi.org/10.1186/1478-7547-2-4>
- Jason, L. A., Benton, M. C., Valentine, L., Johnson, A., & Torres-Harding, S. (2008). The economic impact of ME/CFS: Individual and societal costs. *Dynamic Medicine*, *7*, 6. <https://doi.org/10.1186/1476-5918-7-6>
- Lin, J. S., Resch, S. C., Brimmer, D. J., Johnson, A., Kennedy, S., Burstein, N., & Simon, C. J. (2011). The economic impact of chronic fatigue syndrome in Georgia: Direct and indirect costs. *Cost Effectiveness and Resource Allocation*, *9*(1), 1. <https://doi.org/10.1186/1478-7547-9-1>
- Clayton, E. W. (2015). Beyond myalgic encephalomyelitis/chronic fatigue syndrome: An IOM report on redefining an illness. *JAMA*, *313*(11), 1101–1102. <https://doi.org/10.1001/jama.2015.1346>
- The Centers for Disease Control and Prevention. (2022, March 14). *Myalgic encephalomyelitis/chronic fatigue syndrome*. Retrieved from <https://www.cdc.gov/me-cfs/>
- Chang, C.-J., Hung, L.-Y., Kogelnik, A. M., Kaufman, D., Aiyar, R. S., Chu, A. M., Wilhelmy, J., Li, P., Tannenbaum, L., Xiao, W., & Davis, R. W. (2021). A Comprehensive examination of severely ill ME/CFS patients. *Healthcare*, *9*(10), 1290.
- Friedman, K. J., Murovska, M., Pheby, D. F. H., & Zalewski, P. (2021). Our evolving understanding of ME/CFS. *Medicina (Kaunas, Lithuania)*, *57*(3), 200. <https://doi.org/10.3390/medicina57030200>
- Wong, T. L., & Weitzer, D. J. (2021). Long COVID and myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS)—A systematic review and comparison of clinical presentation and symptomatology. *Medicina (Kaunas, Lithuania)*, *57*(5), 418. <https://doi.org/10.3390/medicina57050418>
- The Centers for Disease Control and Prevention. *Post-COVID Conditions*. Retrieved April 21, 2022 from <https://www.cdc.gov/coronavirus/2019-ncov/long-term-effects/index.html>
- Wanga, V., et al. (2021). Long-term symptoms among adults tested for SARS-CoV-2—United States, January 2020–April 2021. *Morbidity and Mortality Weekly Report (MMWR)*, *70*(36), 1235–1241. <https://doi.org/10.15585/mmwr.mm7036a1>
- Yang, M., Keller, S., & Lin, J.-M.S. (2019). Psychometric properties of the PROMIS® Fatigue Short Form 7a among adults with

- myalgic encephalomyelitis/chronic fatigue syndrome. *Quality of Life Research*, 28(12), 3375–3384.
20. Fukuda, K., Straus, S. E., Hickie, I., Sharpe, M. C., Dobbins, J. G., & Komaroff, A. (1994). The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Annals of Internal Medicine*, 121(12), 953–959.
 21. Carruthers, B. M., Jain, A. K., De Meirleir, K. L., Peterson, D. L., Klimas, N. G., Lerner, A. M., Bested, A. C., Flor-Henry, P., Joshi, P., & Powles, A. P. (2003). Myalgic encephalomyelitis/chronic fatigue syndrome: Clinical working case definition, diagnostic and treatment protocols. *Journal of Chronic Fatigue Syndrome*, 11(1), 7–115.
 22. Carruthers, B. M., van de Sande, M. I., De Meirleir, K. L., Klimas, N. G., Broderick, G., Mitchell, T., Staines, D., Powles, A. C. P., Speight, N., Vallings, R., Bateman, L., Baumgarten-Austrheim, B., Bell, D. S., Carlo-Stella, N., Chia, J., Darragh, A., Jo, D., Lewis, D., Light, A. R., ... Stevens, S. (2011). Myalgic encephalomyelitis: International consensus criteria. *Journal of Internal Medicine*, 270(4), 327–338. <https://doi.org/10.1111/j.1365-2796.2011.02428.x>
 23. Lim, E. J., & Son, C. G. (2020). Review of case definitions for myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). *Journal of Translational Medicine*, 18(1), 289. <https://doi.org/10.1186/s12967-020-02455-0>
 24. Jackson, M. L., & Bruck, D. (2012). Sleep abnormalities in chronic fatigue syndrome/myalgic encephalomyelitis: A review. *Journal of Clinical Sleep Medicine*, 8(6), 719–728. <https://doi.org/10.5664/jcsm.2276>
 25. Mariman, A. N., Vogelaers, D. P., Tobback, E., Delesie, L. M., Hanoulle, I. P., & Pevernagie, D. A. (2013). Sleep in the chronic fatigue syndrome. *Sleep Medicine Reviews*, 17(3), 193–199. <https://doi.org/10.1016/j.smrv.2012.06.003>
 26. Nisenbaum, R., Jones, J. F., Unger, E. R., Reyes, M., & Reeves, W. C. (2003). A population-based study of the clinical course of chronic fatigue syndrome. *Health and Quality of Life Outcomes*, 1, 49–49. <https://doi.org/10.1186/1477-7525-1-49>
 27. Nisenbaum, R., Reyes, M., Unger, E. R., & Reeves, W. C. (2004). Factor analysis of symptoms among subjects with unexplained chronic fatigue: What can we learn about chronic fatigue syndrome? *Journal of Psychosomatic Research*, 56(2), 171–178. [https://doi.org/10.1016/s0022-3999\(03\)00039-4](https://doi.org/10.1016/s0022-3999(03)00039-4)
 28. Hamaguchi, M., Kawahito, Y., Takeda, N., Kato, T., & Kojima, T. (2011). Characteristics of chronic fatigue syndrome in a Japanese community population: Chronic fatigue syndrome in Japan. *Clinical Rheumatology*, 30(7), 895–906. <https://doi.org/10.1007/s10067-011-1702-9>
 29. Watson, N. F., Kapur, V., Arguelles, L. M., Goldberg, J., Schmidt, D. F., Armitage, R., & Buchwald, D. (2003). Comparison of subjective and objective measures of insomnia in monozygotic twins discordant for chronic fatigue syndrome. *Sleep*, 26(3), 324–328. <https://doi.org/10.1093/sleep/26.3.324>
 30. Unger, E. R., Nisenbaum, R., Moldofsky, H., Cesta, A., Sammut, C., Reyes, M., & Reeves, W. C. (2004). Sleep assessment in a population-based study of chronic fatigue syndrome. *BMC Neurology*, 4, 6–6. <https://doi.org/10.1186/1471-2377-4-6>
 31. Creti, L., Libman, E., Baltzan, M., Rizzo, D., Bailes, S., & Fichten, C. S. (2010). Impaired sleep in chronic fatigue syndrome: How is it best measured? *Journal of Health Psychology*, 15(4), 596–607. <https://doi.org/10.1177/1359105309355336>
 32. Maksoud, R., Eaton-Fitch, N., Matula, M., Cabanas, H., Staines, D., & Marshall-Gradisnik, S. (2021). Systematic review of sleep characteristics in myalgic encephalomyelitis/chronic fatigue syndrome. *Healthcare (Basel)*, 9(5), 568. <https://doi.org/10.3390/healthcare9050568>
 33. Watson, N. F., Jacobsen, C., Goldberg, J., Kapur, V., & Buchwald, D. (2004). Subjective and objective sleepiness in monozygotic twins discordant for chronic fatigue syndrome. *Sleep*, 27(5), 973–977. <https://doi.org/10.1093/sleep/27.5.973>
 34. Reeves, W. C., Heim, C., Maloney, E. M., Youngblood, L. S., Unger, E. R., Decker, M. J., Jones, J. F., & Rye, D. B. (2006). Sleep characteristics of persons with chronic fatigue syndrome and non-fatigued controls: Results from a population-based study. *BMC Neurology*, 6, 41. <https://doi.org/10.1186/1471-2377-6-41>
 35. Majer, M., Jones, J. F., Unger, E. R., Youngblood, L. S., Decker, M. J., Gurbaxani, B., Heim, C., & Reeves, W. C. (2007). Perception versus polysomnographic assessment of sleep in CFS and non-fatigued control subjects: Results from a population-based study. *BMC Neurology*, 7, 40. <https://doi.org/10.1186/1471-2377-7-40>
 36. Aaron, L. A., Burke, M. M., & Buchwald, D. (2000). Overlapping conditions among patients with chronic fatigue syndrome, fibromyalgia, and temporomandibular disorder. *Archives of Internal Medicine*, 160(2), 221–227. <https://doi.org/10.1001/archinte.160.2.221>
 37. McGregor, N. R., Armstrong, C. W., Lewis, D. P., Butt, H. L., & Gooley, P. R. (2016). Widespread pain and altered renal function in ME/CFS patients. *Fatigue: Biomedicine, Health & Behavior*, 4(3), 132–145. <https://doi.org/10.1080/21641846.2016.1207400>
 38. Meeus, M., Nijs, J., & Meirleir, K. D. (2007). Chronic musculoskeletal pain in patients with the chronic fatigue syndrome: A systematic review. *European Journal of Pain*, 11(4), 377–386. <https://doi.org/10.1016/j.ejpain.2006.06.005>
 39. Food and Drug Administration. (2013). *The voice of the patient: Chronic fatigue syndrome and myalgic encephalomyelitis*. Center for Drug Evaluation and Research (CDER), Food and Drug Administration.
 40. Unger, E. (2013). *Measures of CFS in a multi-site clinical study*. FDA Scientific Drug Development Workshop.
 41. Aaron, L. A., Arguelles, L. M., Ashton, S., Belcourt, M., Herrell, R., Goldberg, J., Smith, W. R., & Buchwald, D. (2002). Health and functional status of twins with chronic regional and widespread pain. *Journal of Rheumatology*, 29(11), 2426–2434.
 42. Strand, E. B., Castro-Marrero, J., Helland, I., Alegre, J., & Mengshoel, A. M. (2020). Pain and depression are associated with more anxiety in ME/CFS: A cross-sectional cohort study between Norway and Spain. *Clinical Medicine Insights: Psychiatry*, 11, 1–8. <https://doi.org/10.1177/1179557320941478>
 43. Burckhardt, C. S., Clark, S. R., & Bennett, R. M. (1991). The fibromyalgia impact questionnaire: Development and validation. *Journal of Rheumatology*, 18(5), 728–733.
 44. Bennett, R. M., Friend, R., Jones, K. D., Ward, R., Han, B. K., & Ross, R. L. (2009). The Revised Fibromyalgia Impact Questionnaire (FIQR): Validation and psychometric properties. *Arthritis Research & Therapy*, 11(4), R120. <https://doi.org/10.1186/ar2783>
 45. Cleeland, C. S., Gonin, R., Hatfield, A. K., Edmonson, J. H., Blum, R. H., Stewart, J. A., & Pandya, K. J. (1994). Pain and its treatment in outpatients with metastatic cancer. *New England Journal of Medicine*, 330(9), 592–596. <https://doi.org/10.1056/nejm199403033300902>
 46. Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care*, 30(6), 473–483.
 47. Jason, L. A., Evans, M., Porter, N., Brown, M., Brown, A., Hunnell, J., Anderson, V., Lerch, A., Meirleir, K. D., & Friedberg, F. (2010). The development of a revised Canadian myalgic encephalomyelitis chronic fatigue syndrome case definition. *American Journal of Biochemistry and Biotechnology*, 6(2), 120–135. <https://doi.org/10.3844/ajbb.2010.120.135>

48. Wagner, D., Nisenbaum, R., Heim, C., Jones, J. F., Unger, E. R., & Reeves, W. C. (2005). Psychometric properties of the CDC Symptom Inventory for assessment of chronic fatigue syndrome. *Population Health Metrics*, 3, 8. <https://doi.org/10.1186/1478-7954-3-8>
49. Jason, L. A., Ropacki, M. T., Santoro, N. B., Richman, J. A., Heatherly, W., Taylor, R., Ferrari, J. R., Haney-Davis, T. M., Rademaker, A., Dupuis, J., Golding, J., Plioplys, A. V., & Plioplys, S. (1997). A screening instrument for chronic fatigue syndrome: Reliability and validity. *Journal of Chronic Fatigue Syndrome*, 3(1), 39–59. https://doi.org/10.1300/J092v03n01_04
50. Bjorner, J. B., Rose, M., Gandek, B., Stone, A. A., Junghaenel, D. U., & Ware, J. E., Jr. (2014). Method of administration of PROMIS scales did not significantly impact score level, reliability, or validity. *Journal of Clinical Epidemiology*, 67(1), 108–113. <https://doi.org/10.1016/j.jclinepi.2013.07.016>
51. Embretson, S. E., & Reise, S. P. (2013). *Item response theory*. Psychology Press.
52. Liu, H., Cella, D., Gershon, R., Shen, J., Morales, L. S., Riley, W., & Hays, R. D. (2010). Representativeness of the PROMIS internet panel. *Journal of Clinical Epidemiology*, 63(11), 1169–1178. <https://doi.org/10.1016/j.jclinepi.2009.11.021>
53. Yu, L., Buysse, D. J., Germain, A., Moul, D. E., Stover, A., Dodds, N. E., Johnston, K. L., & Pilkonis, P. A. (2011). Development of short forms from the PROMIS™ sleep disturbance and Sleep-Related Impairment item banks. *Behavioral Sleep Medicine*, 10(1), 6–24. <https://doi.org/10.1080/15402002.2012.636266>
54. Amtmann, D., Cook, K. F., Jensen, M. P., Chen, W. H., Choi, S., Revicki, D., Cella, D., Rothrock, N., Keefe, F., Callahan, L., & Lai, J. S. (2010). Development of a PROMIS item bank to measure pain interference. *Pain*, 150(1), 173–182. <https://doi.org/10.1016/j.pain.2010.04.025>
55. Revicki, D. A., Chen, W.-H., Harnam, N., Cook, K. F., Amtmann, D., Callahan, L. F., Jensen, M. P., & Keefe, F. J. (2009). Development and psychometric analysis of the PROMIS pain behavior item bank. *Pain*, 146(1), 158–169. <https://doi.org/10.1016/j.pain.2009.07.029>
56. SAS Institute Inc. (2002–2014). *SAS® 9.4 help and documentation*. SAS Institute Inc.
57. Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide. Eighth Edition*. Muthén and Muthén.
58. Cai, L., Thissen, D., & du Toit, S. H. C. (2011). *IRTPRO: Flexible, multidimensional, multiple categorical IRT modeling*. Scientific Software International.
59. Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34–42.
60. Wang, L., Zhang, Z., McArdle, J. J., & Salthouse, T. A. (2009). Investigating ceiling effects in longitudinal data analysis. *Multivariate Behavioral Research*, 43(3), 476–496. <https://doi.org/10.1080/00273170802285941>
61. McDonald, R. P. (1999). *Test theory: A unified treatment*. Erlbaum.
62. Reise, S. P., Morizot, J., & Hays, R. D. (2007). The role of the bifactor model in resolving dimensionality issues in health outcomes measures. *Quality of Life Research*, 16(1), 19–31.
63. Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, 21(2), 137–150.
64. Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
65. Samejima, F. (1997). Graded response model. *Handbook of modern item response theory* (pp. 85–100). Springer.
66. Green, S. B., & Yang, Y. (2009). Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika*, 74(1), 155–167. <https://doi.org/10.1007/s11336-008-9099-3>
67. Yang, Y., & Xia, Y. (2019). Categorical omega with small sample sizes via Bayesian estimation: An alternative to frequentist estimators. *Educational and Psychological Measurement*, 79(1), 19–39. <https://doi.org/10.1177/0013164417752008>
68. Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika*, 74(1), 137–143.
69. Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399–412.
70. Peters, G.-J. (2014). The alpha and the omega of scale reliability and validity: Why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *European Health Psychologist*, 16(2), 56–69. <https://doi.org/10.31234/osf.io/h47fv>
71. Nunnally, J. C. (1978). *Psychometric theory*. McGraw-Hill.
72. Lohr, K. N. (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research*, 11(3), 193–205.
73. Bartlett, S. J., Orbai, A. M., Duncan, T., DeLeon, E., Ruffing, V., Clegg-Smith, K., & Bingham, C. O., III. (2015). Reliability and validity of selected PROMIS measures in people with rheumatoid arthritis. *PLoS ONE*, 10(9), e0138543. <https://doi.org/10.1371/journal.pone.0138543>
74. Quach, C. W., Langer, M. M., Chen, R. C., Thissen, D., Usinger, D. S., Emerson, M. A., & Reeve, B. B. (2016). Reliability and validity of PROMIS measures administered by telephone interview in a longitudinal localized prostate cancer study. *Quality of Life Research*, 25(11), 2811–2823. <https://doi.org/10.1007/s11136-016-1325-3>
75. Merriwether, E. N., Rakel, B. A., Zimmerman, M. B., Dailey, D. L., Vance, C. G. T., Darghosian, L., Golchha, M., Geasland, K. M., Chimenti, R., Crofford, L. J., & Sluka, K. A. (2017). Reliability and construct validity of the patient-reported outcomes measurement information system (PROMIS) instruments in women with fibromyalgia. *Pain Medicine (Malden, Mass.)*, 18(8), 1485–1495. <https://doi.org/10.1093/pm/pnw187>
76. Full, K. M., Malhotra, A., Crist, K., Moran, K., & Kerr, J. (2019). Assessing psychometric properties of the PROMIS Sleep Disturbance Scale in older adults in independent-living and continuing care retirement communities. *Sleep Health*, 5(1), 18–22. <https://doi.org/10.1016/j.sleh.2018.09.003>
77. Chimenti, R. L., Rakel, B. A., Dailey, D. L., Vance, C. G. T., Zimmerman, M. B., Geasland, K. M., Williams, J. M., Crofford, L. J., & Sluka, K. A. (2021). Test-retest reliability and responsiveness of PROMIS sleep short forms within an RCT in women with fibromyalgia. *Frontiers in Pain Research*, 2, 682072. <https://doi.org/10.3389/fpain.2021.682072>
78. Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Routledge.
79. Langer, M. M. (2008). *A reexamination of Lord's Wald test for differential item functioning using item response theory and modern error estimation* [Doctoral Dissertation, University of North Carolina at Chapel Hill]. Retrieved from <https://cdr.lib.unc.edu/record/uuid:3ef47e17-9b76-45be-bf49-febddd17f4e0>
80. Reyes, M., Nisenbaum, R., Hoaglin, D. C., Unger, E. R., Emmons, C., Randall, B., Stewart, J. A., Abbey, S., Jones, J. F., Gantz, N., & Minden, S. (2003). Prevalence and incidence of chronic fatigue syndrome in Wichita, Kansas. *Archives of Internal Medicine*, 163(13), 1530–1536. <https://doi.org/10.1001/archinte.163.13.1530>

81. Rusu, C., Gee, M. E., Lagacé, C., & Parlor, M. (2015). Chronic fatigue syndrome and fibromyalgia in Canada: Prevalence and associations with six health status indicators. *Health Promotion and Chronic Disease Prevention in Canada: Research, Policy and Practice*, 35(1), 3–11.
82. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300.
83. Thissen, D., Steinberg, L., & Kuang, D. (2002). Quick and easy implementation of the Benjamini-Hochberg procedure for controlling the false positive rate in multiple comparisons. *Journal of Educational and Behavioral Statistics*, 27(1), 77–83. <https://doi.org/10.3102/10769986027001077>
84. National Institute of Neurological Disorders and Stroke. (2022, March 14). *Common data elements for myalgic encephalomyelitis/chronic fatigue syndrome*. Retrieved from <https://www.commodataelements.ninds.nih.gov/Myalgic%20Encephalomyelitis/Chronic%20Fatigue%20Syndrome>
85. Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 5(2), 99–114. <https://doi.org/10.2307/3001913>
86. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.
87. Miles, J., & Shevlin, M. (2001). *Applying regression & correlation: A guide for students and researchers* (1st ed.). Sage Publications.
88. Badhiwala, J. H., Witiw, C. D., Nassiri, F., Akbar, M. A., Jaja, B., Wilson, J. R., & Fehlings, M. G. (2018). Minimum clinically important difference in SF-36 scores for use in degenerative cervical myelopathy. *Spine*. <https://doi.org/10.1097/brs.00000000000002684>
89. Swigris, J. J., Brown, K. K., Behr, J., du Bois, R. M., King, T. E., Raghu, G., & Wamboldt, F. S. (2010). The SF-36 and SGRQ: Validity and first look at minimum important differences in IPF. *Respiratory medicine*, 104(2), 296–304. <https://doi.org/10.1016/j.rmed.2009.09.006>
90. Ward, M. M., Guthrie, L. C., & Alba, M. I. (2014). Clinically important changes in short form-36 scales for use in rheumatoid arthritis clinical trials: The impact of low responsiveness. *Arthritis Care & Research*, 66(12), 1783–1789. <https://doi.org/10.1002/acr.22392>
91. Norman, G. R., Sloan, J. A., & Wywich, K. W. (2003). Interpretation of changes in health-related quality of life: The Remarkable Universality of half a standard deviation. *Medical Care*, 41(5), 582–592. <https://doi.org/10.1097/01.mlr.0000062554.74615.4c>
92. Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases*, 40(2), 171–178. [https://doi.org/10.1016/0021-9681\(87\)90069-5](https://doi.org/10.1016/0021-9681(87)90069-5)
93. Cook, K. F., Kallen, M. A., & Amtmann, D. (2009). Having a fit: Impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. *Quality of Life Research*, 18(4), 447–460. <https://doi.org/10.1007/s11136-009-9464-4>
94. Hildenbrand, A. K., Quinn, C. T., Mara, C. A., Peugh, J. L., McTate, E. A., Britto, M. T., & Crosby, L. E. (2019). A preliminary investigation of the psychometric properties of PROMIS® scales in emerging adults with sickle cell disease. *Health Psychology*, 38(5), 386–390. <https://doi.org/10.1037/hea0000696>
95. Becker, B., Raymond, K., Hawkes, C., Foster, A. M., Lovley, A., Saucier, C., Rizio, A. A., Bjorner, J. B., & Kosinski, M. (2021). Qualitative and psychometric approaches to evaluate the PROMIS pain interference and sleep disturbance item banks for use in patients with rheumatoid arthritis. *Journal of Patient-Reported Outcomes*, 5(1), 52. <https://doi.org/10.1186/s41687-021-00318-w>
96. Yost, K. J., Eton, D. T., Garcia, S. F., & Cella, D. (2011). Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. *Journal of Clinical Epidemiology*, 64(5), 507–516. <https://doi.org/10.1016/j.jclinepi.2010.11.018>
97. Chen, C. X., Kroenke, K., Stump, T. E., Kean, J., Carpenter, J. S., Krebs, E. E., Bair, M. J., Damush, T. M., & Monahan, P. O. (2018). Estimating minimally important differences for the PROMIS pain interference scales: Results from 3 randomized clinical trials. *Pain*, 159(4), 775–782. <https://doi.org/10.1097/j.pain.0000000000001121>
98. Purvis, T. E., Neuman, B. J., Riley, L. H., 3rd., & Skolasky, R. L. (2018). Discriminant ability, concurrent validity, and responsiveness of PROMIS health domains among patients with lumbar degenerative disease undergoing decompression with or without arthrodesis. *Spine*, 43(21), 1512–1520. <https://doi.org/10.1097/brs.0000000000002661>
99. Katz, P., Pedro, S., Alemao, E., Yazdany, J., Dall'Era, M., Trupin, L., Rush, S., & Michaud, K. (2020). Estimates of responsiveness, minimally important differences, and patient acceptable symptom state in five patient-reported outcomes measurement information system short forms in systemic lupus erythematosus. *ACR Open Rheumatology*, 2(1), 53–60. <https://doi.org/10.1002/acr2.11100>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.