Clinical Validity of the PROMIS® Fatigue Item Bank across Diverse Clinical Samples

David Cella[1], **Jin-Shei Lai**[1], **Sally E. Jensen**[1], **Christopher Christodoulou**[2], **Doerte U. Junghaenel**[3], **Bryce B. Reeve**[4], and **Arthur A. Stone**[3]

[1]Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

[2]Department of Neurology, Stony Brook University, Stony Brook, NY, USA

[3]University of Southern California, Los Angeles, CA, USA

[4]Department of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC

Abstract

Objective—To evaluate the comparability and responsiveness of PROMIS Fatigue Item Bank across six chronic conditions.

Study Design and Setting—Individuals (n=1,430) with chronic obstructive pulmonary disease (COPD; n=125), chronic heart failure (CHF; n=60), chronic back pain (n=218), major depressive disorder (MDD; n=196), rheumatoid arthritis (RA; n=521), and cancer(n=310) completed assessments from the PROMIS fatigue item bank at baseline and a clinically-relevant follow-up. The cancer and arthritis samples were followed in observational studies; the other four groups were enrolled immediately prior to a planned clinical intervention. All participants completed global ratings of change at follow-up. Linear mixed effects models and standardized response means were estimated to examine clinical validity and responsiveness to change.

Results—All patient groups reported more fatigue than the general population (range = 0.2 – 1.29 SD worse). The four clinical groups with pre-treatment baseline data experienced significant

improvement in fatigue at follow-up (effect size range= 0.25 to 0.91). Individuals reporting better overall health usually experienced larger fatigue changes than those reporting worse overall health.

Conclusion—The results support the PROMIS fatigue measures's responsiveness to change in six different chronic conditions. In addition, these results support the ability of the PROMIS fatigue measures to compare differences in fatigue across a range of chronic conditions, thereby enabling comparative effectiveness research.

Keywords

PROMIS; Fatigue; Chronic Conditions; Item Bank

1. Introduction

Fatigue is a symptom commonly experienced by healthy individuals as well as those with chronic disorders. When experienced as part of a chronic condition, it is often experienced as overwhelming, debilitating, and exhausting; decreasing one's ability to carry out daily activities, including the ability to work effectively and to function at one's usual level in family or social roles.[1–4] A growing body of literature documents the high prevalence of fatigue and its impact across a variety of chronic health conditions, including back pain,[5] cancer,[6, 7] congestive heart failure (CHF),[8] chronic obstructive pulmonary disease (COPD),[9] major depressive disorder (MDD),[10, 11] and rheumatoid arthritis (RA).[12] Its prevalence makes it a common treatment target, as relieving fatigue often results in improved well-being and function across a large number of people.

The World Health Organization (WHO)'s International Classification of Functioning, Disability and Health included the minimization of fatigue among its stated aims,[13] highlighting the importance of regular assessment of fatigue in both research and clinical contexts. Although a number of disease-specific fatigue measures exist,[14–17] a well-developed and carefully-calibrated universal fatigue measure that can be applied across chronic health populations and treatment contexts could enhance the comparability of findings and thus serve as a common metric of fatigue across chronic health condition groups. This would greatly enhance the interpretability of fatigue results across clinical research studies, and enable meaningful comparative effectiveness research.

To this end, the Patient-Reported Outcomes Measurement Information System® (PROMIS®) investigators utilized a multi-step, mixed methods approach to develop a fatigue item bank which can be used as an assessment tool as either a computerized adaptive test (CAT) or a fixed-length short form. The development process and psychometric properties of the fatigue item bank have been reported previously.[18–20] In this paper, we describe the longitudinal clinical validation of the fatigue item bank in adults in six different chronic health condition samples: back pain, cancer, CHF, COPD, MDD, and RA. We hypothesized that these clinical samples would present with more fatigue than is found in the general US population. We also hypothesized that clinical samples with baseline (pre-treatment) data available, who were embarking on a new or modified treatment plan (i.e., treatment for back pain; CHF; COPD patients in an acute exacerbation of symptoms; MDD), would experience longitudinal improvements in fatigue. We also predicted relative stability in fatigue scores

over time among other clinical samples being followed naturalistically (i.e., COPD-stable patients; cancer; RA), or in clinical stituations where some patients would be expected to improve, some would be expected to worsen, and many would not be expected to change (e.g., RA, stable COPD). Finally, we hypothesized that fatigue scores would differentiate subsets of samples that were distinct in terms of clinical severity or functional impairment (e.g., COPD-stable versus COPD-exacerbation).

2. Methods

2.1. Clinical samples

Included in this report are data collected across six studies (back pain, cancer, CHF, COPD, MDD, and RA), conducted by PROMIS investigators. The studies of MDD, back pain, and CHF followed patients as they enrolled in new treatments. Patients with acute COPD exacerbation were expected to experience symptom resolution over the course of the study. Both RA and cancer samples were heterogeneous with respect to intervention, but were dominated by participants who were already receiving treatments by the time they enrolled in the current study. We examined the longitudinal data at baseline and follow-up, namely, 3 months after start of study (MDD, back pain, and COPD), 8–12 weeks after heart transplantation (CHF), 6–12 weeks after enrollment (cancer), and 12 months after enrollment (RA). Although the COPD-stable, cancer, and RA groups were not enrolled in new treatments, we apply the clinical trial terms "baseline" and "follow-up" to all study groups for consistency. Details of the sample information and recruitment procedures are described in Cook et al.[21] (this volume).

2.2. PROMIS Fatigue item bank

The PROMIS Fatigue item bank is comprised of 95 items, including the 13-item FACIT-Fatigue,[14] calibrated from an initial pool of 112 items tapping two conceptually related areas: fatigue experience and fatigue interference in daily life and function.[18] Higher scores suggest worse fatigue. CAT and short-forms derived from the item bank can reliably estimate fatigue, with scores referenced to the US general population using the T-score metric, with mean=50 and standard deviation (SD)=10. Fatigue T-scores were estimated via CAT for patients enrolled in the CHF, COPD, back pain, and MDD studies. In cancer and RA, Fatigue T-scores were obtained using the PROMIS Fatigue short-form version 1. Since both short-form and CAT used the same item parameters as used by the PROMIS Fatigue item bank, their resulting fatigue scores are comparable.[18] PROMIS Fatigue items are available in an online Appendix.

2.2. Statistical analyses

Analyses to evaluate responsiveness to change were conducted separately for each of seven clinical groups: back pain, cancer, CHF, COPD-exacerbation, COPD-stable, MDD, and RA. Items measuring general health (e.g., In general, would you say your health is: Excellent, Very Good, Good, Fair, or Poor) were used to estimate patient-perceived responsiveness for each condition. For these items, the baseline and follow-up scores were subtracted to find change scores. In addition, fatigue specific global change items were used to evaluate responsiveness for cancer and RA (e.g., "Since the last time you filled out a questionnaire,

your level of fatigue is: very much better, moderately better, a little better, about the same, a little worse, moderately worse, or very much worse"). For change scores of general health items and fatigue global change items, scores were grouped into three categories - better, about the same, and worse – for the responsiveness analyses.

Linear mixed models were estimated with random subject effects to account for the similarity among repeated observations within individuals.[22, 23] Missing data were evaluated prior to performing longitudinal analyses. Since it was reasonable to consider the missing data to be missing completely at random (MCAR) or missing at random (MAR), a mixed model is advantageous because all available data can be used; in other words, the analyses were not restricted to only those respondents with data at both time points.[24, 25] Least squares means, standard errors and 95% confidence intervals were estimated from the models.

Change scores in the PROMIS Fatigue T-scores (both from CAT and SF) were used to estimate the standardized response mean (SRM) for each of the three change groups. This is the ratio of the mean change to the standard deviation of that change.[26] It is a form of Cohen's effect size index.[27] We set an SRM of 0.30 as the minimum required magnitude for difference or change scores, to consider them as candidates for clinical meaningfulness. [28, 29]

3. Results

3.1. Participant characteristics

A total of 1,430 people participated in the clinical studies. Participants were diverse in terms of gender, age and marital status, as reported in the overview paper in this issue (this volume).[21] Most participants were non-Hispanic White, had some college education, and had moderate to severe health limitations. At baseline, participants from all studies reported more fatigue, ranging from 2–13 T-score points (i.e.,0.2 – 1.3 SD on the T-score metric) higher than the US population norm (score of 50). Significantly different fatigue scores were reported across these conditions, F=34.97, p<.0001. Specifically, COPD-exacerbation group reported significantly (p<0.05) more fatigue than back pain, COPD-stable, RA and Cancer. MDD group reported significantly (p<0.05) more fatigue than back pain, COPD-stable, RA, and Cancer. CHF group reported significantly (p<0.05) more fatigue than RA and Cancer. Fatigue T-scores of each condition were 56.7 (SD=9.4) for back pain, 52.0 (SD=7.6) for cancer, 58.9 (SD=10.4) for CHF, 62.8 (SD=8.3) for COPD-exacerbation, 56.3 (SD=8.6) for COPD-stable, 61.3 (SD=8.3) for MDD, and 53.8 (SD=8.8) for RA (see Table 1a).

3.2. Responsiveness

Data from those who completed the follow-up assessments (87.6%; n=1,252) were used for responsiveness analyses. All four diagnostic groups that were enrolled in single-arm intervention studies (COPD; CHF; MDD; back pain) experienced predicted significant improvement in fatigue over time. The two groups enrolled in observational studies (cancer; RA) did not have group-wide changes in their fatigue levels over time. Group-wide fatigue change were not predicted in these two cohorts. Rather, we divided these two groups into

improved and worsened subgroups based on patient-reported global ratings of change in fatigue at follow-up. The results of the mixed models are summarized in Tables 1a and 1b. Least squares means from baseline to time 2 are shown in Figure 1. Estimated mean change scores for the improving groups ranged from 2.4 (COPD-Stable) to 11.6 (CHF). Negligible change scores were found on RA (0.40) and slightly more fatigue was found on Cancer (1.16 points) when they were evaluated as a whole. Yet, when evaluating both groups by subgroups, change scores for improved groups were −0.27 and −1.17, while change scores for worsened groups were 1.44 and 4.67 for RA and Cancer, respectively.

For most participants who reported that their overall health and fatigue (for Cancer, CHF, COPD-exacerbation, COPD-stable and RA) changed for the better, there was a corresponding improvement in fatigue over time with T-score changes ranging from −0.35 (RA) to −11.9 (CHF) (standardized response means ranged from −0.06 to −1.23; see Table 2). Effect sizes (i.e., SRM) for those who reported better overall health (i.e., global health) or fatigue at follow up were always larger than effect sizes for those who reported worse overall health except for Cancer (both overall health and fatigue) and RA (fatigue only). Yet, mixed responses were found for those who reported their overall health and fatigue *worsened* or were *about the same*. Only the change scores of the RA patients and cancer patients showed significantly worse fatigue scores at follow-up (Table 2).

4. Discussion

The PROMIS Fatigue item bank was developed using rigorous methods, demonstrates good psychometric properties, and is publicly available.[18] The present study extends the intial published information on reliability and validity by examining the longitudinal validation of the PROMIS Fatigue item bank in six chronic conditions, thus providing support for the clinical validity of PROMIS fatigue measures. The findings also highlight the ability to meaningfully compare fatigue levels across chronic disease samples, providing an evidence base to support the setting of responder definitions and to enable comparative effectiveness research that relies upon cross-disease comparisons, or within-disease comparisons across treatments.

Responsiveness, or sensitivity to detect change in fatigue over time and in response to clinical intervention, represents an important attribute for fatigue PROs and remains essential for their acceptance in clinical research and practice. Moreover, the ability to detect bidirectional fatigue change in terms of improvement and deterioration, while acknowledging score variability unrelated to change (i.e., error), constitute important characteristics of fatigue measures. In the present study, PROMIS fatigue scores improved over time in all four groups that were enrolled into single-arm trials that were designed to detect clinically-anticipated improvement at the group level. This included CHF, back pain, MDD, and COPD-exacerbation samples. This is consistent with our hypothesis that clinical samples undergoing condition-targeted interventions would report a post-treatment decrease in fatigue relative to baseline. The improvement in PROMIS fatigue noted in the COPD-stable subgroup was not expected. It is possible that this subgroup improved slightly due to change in management initiated at the baseline visit, or that the improved score (although lower in magnitude than that observed in the other clinical groups, including the COPD

exacerbation group) was a random (chance) observation. The low magnitude of improvement (2.4 units) relative to the others (range = 3.5 – 11.6 units) is of a magnitude that may not be clinically meaningful. For example, Yost et al estimated the minimally important difference of two PROMIS fatigue instruments to be in the range of 2.5–5.0 points.[30] By contrast, the very large fatigue improvement in CHF patients (Table 2; SRM >1.0) might reflect a uniquely dramatic benefit of a major surgical intervention for this condition. Further research can evaluate the magnitude of benefit of other surgical and medical treatment options for CHF and associated fatigue.

The responsiveness analyses for arthritis and cancer samples centered around the patients' global rating of change. In both cases, and in both directions of change (improvement and worsening), we observed changes in the PROMIS fatigue scores in the predicted direction. These changes were statistically significant for those patients who said they were *worse* on the fatigue-specific global question, but not for those patients who said they were *better*. This asymmetry (i.e., relative to worsening change scores, smaller improvement change scores are considered meaningful to patients) has been noted previously.[31] Interestingly, PROMIS fatigue change scores associated with patients who said they felt generally better or generally worse were also consistently in the predicted direction, and in the case of people who reported feeling globally better, were of a higher magnitude than the fatigue-specific global rating (Table 2).

These findings provide evidence to suggest that the PROMIS fatigue measures used in these studies are clinically valid and allow for direct comparisons across six common chronic conditions. Regarding responsiveness to change, the PROMIS fatigue measures used across these studies detected upward and downward change in most cases. Moreover, the findings suggest that the PROMIS fatigue measures are robust to non-change related variability among clinical samples whose conditions would not be expected to change substantially over time, except COPD-stable as discussed above.

The ability for PROMIS fatigue measures to differentiate between subgroups of a clinical sample that differ in severity constitutes another important characteristic when evaluating clinical validity. The present study provided the opportunity to examine this question by comparing fatigue scores between the COPD-stable and COPD-exacerbation group. As hypothesized, the COPD-exacerbation subgroup reported significantly greater fatigue than their COPD-stable counterparts at both baseline and follow-up. We also predicted that these two groups would differ in longitudinal fatigue changes, with expected improvement in fatigue scores for the COPD-exacerbation group, but not the COPD-stable group. Although the findings demonstrated unexpected improvement in fatigue over time in this COPD-stable group, it should be noted that the magnitude of fatigue improvement in the COPD-stable group was smaller than the magnitude of improvement among the COPD-exacerbation, thus reflecting relatively greater stability of fatigue over time in the COPD-stable group when compared to the COPD-exacerbation. This pattern of findings indicates that the PROMIS fatigue measure is sensitive to detect differences between subgroups of a clinical sample with varying severity levels.

When considering the results of this study, several limitations should be noted. First, none of the four planned intervention studies included a control group. As a result, we cannot differentiate the intervention effect from a placebo effect; such a determination can only be made with a randomized controlled trial. Small sample sizes (n<5) in some cells might not provide enough statistical power which may have affected the interpretation of results. Because of their inherently uncontrolled nature, observational studies are not ideal for evaluating responsiveness of outcome measures such as those evaluated here. Second, the use of patient global ratings of change, while face valid and clinically-relevant, comes with some problems of methodology and interpretation. Because they are gathered at follow-up, global ratings of change are typically more highly-correlated with post-test than they are with pretest, or with the change score itself.[32] Nevertheless, they provide a useful starting "anchor" for estimating the magnitude of measured change that is important to patients at follow-up assessment.

Despite these limitations, the present study extends previous validation of the PROMIS fatigue measure by examining longitudinal change in scores across a diverse set of clinical samples. This provided the opportunity to assess the clinical validity of this measure by examining responsiveness to change following a treatment intervention, examining stability of fatigue scores over time in stable clinical conditions not receiving an intervention, and by examining differences in fatigue scores between different severity level subgroups of the same chronic condition. Although full clinical validity remains a dynamic construct which is never fully achieved, but rather continuously examined, the present study provides an important step in facilitating the productive application of the PROMIS fatigue measures in clinical care and research, most particularly comparative effectiveness research.

Acknowledgments

References

1. Glaus, A. Fatigue in patients with cancer: Analysis and assessment. Heidelberg, Germany: Springer-Verlag Berlin; 1998.

2. North American Nursing Diagnosis Association. Nursing diagnoses: Definition and Classification, 1997–1998. Philadelphia, PA: McGraw-Hill; 1996.

3. Stewart, AL.; Hays, R.; Ware, JE. Health perceptions, energy/fatigue, and health distress measures. In: Stewart, AL.; Ware, JE., editors. Measuring functioning and well-being : the medical outcomes study approach. Durham: Duke University Press; 1992. p. 143-172.

4. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Amtmann D, Bode R, Buysse D, Choi S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. J Clin Epidemiol. 2010; 63(11):1179–1194. [PubMed: 20685078]

5. Salvetti MG, Pimenta CA, Braga PE, M M. Prevalence of fatigue and associated factors in chronic low back pain patients. Rev Lat Am Enfermagem. 2013; 21(Spec No)

6. Berger AM, Abernethy AP, Atkinson A, Barsevick AM, Breitbart WS, Cella D, Cimprich B, Cleeland C, Eisenberger MA, Escalante CP, et al. Cancer-Related Fatigue. J Natl Compr Canc Netw. 2010; 8(8):904–931. [PubMed: 20870636]

7. Piper BF, Cella D. Cancer-Related Fatigue: Definitions and Clinical Subtypes. J Natl Compr Canc Netw. 2010; 8(8):958–966. [PubMed: 20870639]

8. Falk K, Patel H, Swedberg K, Ekman I. Fatigue in Patients with Chronic Heart Failure — A Burden Associated with Emotional and Symptom Distress. Eur J Cardiovasc Nurs. 2009; 8(2):91–96. [PubMed: 18715830]

9. Stridsman C, Müllerova H, Skär L, Lindberg A. Fatigue in COPD and the Impact of Respiratory Symptoms and Heart Disease—A Population-based Study. COPD. 2013; 10(2):125–132. [PubMed: 23547627]

10. Schrepf A, Clevenger L, Christensen D, DeGeest K, Bender D, Ahmed A, Goodheart MJ, Dahmoush L, Penedo F, Lucci JA Iii, et al. Cortisol and inflammatory processes in ovarian cancer patients following primary treatment: Relationships with depression, fatigue, and disability. Brain Behav Immun. 2012 Epub ahead of print.

11. Marcus SM, Young EA, Kerber KB, Kornstein S, Farabaugh AH, Mitchell J, Wisniewski SR, Balasubramani GK, Trivedi MH, Rush AJ. Gender differences in depression: Findings from the STAR*D study. J. Affect. Disord. 2005; 87(2–3):141–150. [PubMed: 15982748]

12. Pollard LC, Choy EH, Gonzalez J, Khoshaba B, Scott DL. Fatigue in rheumatoid arthritis reflects pain, not disease activity. Rheumatology. 2006; 45(7):885–889. [PubMed: 16449363]

13. World Health Organization. Geneva, Switzerland: World Health Organization; 2001. International Classification of Functioning, Disability and Health (ICF).

14. Yellen SB, Cella DF, Webster K, Blendowski C, Kaplan E. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. J Pain Symptom Manage. 1997; 13(2):63–74. [PubMed: 9095563]

15. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, De Haes JC. The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993; 85(5):365–376. [PubMed: 8433390]

16. Piper BF, Dibble SL, Dodd MJ, Weiss MC, Slaughter RE, Paul SM. The revised Piper Fatigue Scale: Psychometric Evaluation in Women with Breast Cancer. Oncol Nurs Forum. 1998; 25(4):677–684. [PubMed: 9599351]

17. Stein KD, Martin SC, Hann DM, Jacobsen PB. A multidimensional measure of fatigue for use with cancer patients. Cancer Pract. 1998; 6(3):143–152. [PubMed: 9652245]

18. Lai JS, Cella D, Choi SW, Junghaenel DU, Christodolou C, Gershon R, Stone A. How Item Banks and Their Application Can Influence Measurement Practice in Rehabilitation Medicine: A PROMIS Fatigue Item Bank Example. Arch Phys Med Rehabil. 2011; 92(10 Supplement):S20–S27. [PubMed: 21958919]

19. Christodoulou C, Junghaenel DU, DeWalt DA, Rothrock N, Stone AA. Cognitive interviewing in the evaluation of fatigue items: Results from the patient-reported outcomes measurement information system (PROMIS). Qual Life Res. 2008; 17(10):1239–1246. [PubMed: 18850327]

20. Junghaenel DU, Christodoulou C, Lai JS, Stone AA. Demographic correlates of fatigue in the US general population: results from the patient-reported outcomes measurement information system (PROMIS) initiative. J Psychosom Res. 2011; 71(3):117–123. [PubMed: 21843744]

21. Cook K, Jensen SE, Schalet BD, Beaumont JL, Amtmann D, Czajkowski S, Dewalt D, Fries JF, Johnson LL, Pilkonis P, et al. PROMIS measures of pain, fatigue, negative affect, physical function and social function demonstrate clinical validity across a range of chronic conditions. J Clin Epidemiol Submitted.

22. Verbeke, G.; Molenberghs, G. Linear Mixed Models for Longitudinal Data. New York, NY: Springer-Verlag; 2000.

23. Hedeker, DR.; Gibbons, RD. Longitudinal data analysis. Hoboken, N.J.: Wiley-Interscience; 2006.

24. Troxel AB, Fairclough DL, Curran D, Hahn EA. Statistical analysis of quality of life with missing data in cancer clinical trials. Stat Med. 1998; 17(5–7):653–666. [PubMed: 9549814]

25. Little, RJA.; Rubin, DB. Statistical Analysis with Missing Data. Hoboken, NJ: John Wiley & Sons, Inc.; 2002.

26. Walker CM, Beretvas SN. Comparing Multidimensional and Unidimensional Proficiency Classifications: Multidimensional IRT as a Diagnostic Aid. Journal of Educational Measurement. 2003; 40(3):255–275.

27. Cohen, J. Statistical power analysis for the behavioral sciences. 2nd. Hillsdale, N.J.: L. Erlbaum Associates; 1988.

28. Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: The FACIT experience. Eval Health Prof. 2005; 28(2):172–191. [PubMed: 15851772]

29. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. Health Qual Life Outcomes. 2006; 4:70. [PubMed: 17005038]

30. Yost KJ, Eton DT, Garcia SF, Cella D. Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. J Clin Epidemiol. 2011; 64(5):507–516. [PubMed: 21447427]

31. Cella D, Hahn EA, Dineen K. Meaningful change in Cancer-Specific Quality-of-Life Scores: Differences Between Improved and Worsening. Qual Life Res. 2002; 11(3):207–221. [PubMed: 12074259]

32. Norman G, Stratford P, Regehr G. Bias in retrospective calculation of responsiveness to change. Journal of Clinical Epidemiology. 1997; 8:869–879. [PubMed: 9291871]

33. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care. 1990; 28(7):632–642. [PubMed: 2366602]

34. Liu H, Cella D, Gershon R, Shen J, Morales LS, Riley W, Hays RD. Representativeness of the Patient-Reported Outcomes Measurement Information System Internet Panel. J Clin Epidemiol. 2010; 63(11):1169–1178. [PubMed: 20688473]
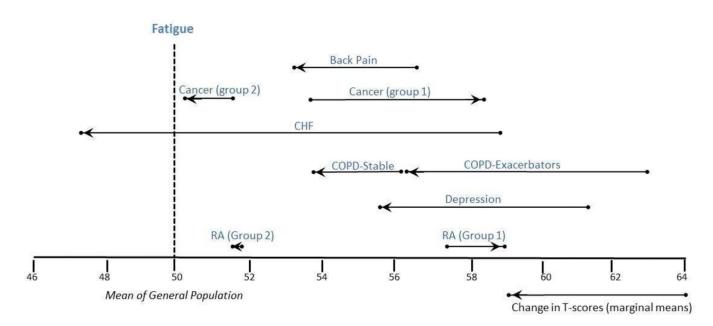
What is new?

Key Findings

- PROMIS Fatigue item bank is a valid tool to measure fatigue experienced by people with diverse chronic conditions.

What this adds to what was known?

- PROMIS Fatigue measures can detect change over time in people with a range of chronic conditions.

What is the implication and what should change now?

- These results provide an important step in facilitating the use and acceptance of the PROMIS fatigue measures in clinical practice and comparative effectiveness research.

- PROMIS Fatigue measures are publicly available (www.assessmentcenter.net). Custom short forms can be designed and scored using PROMIS item response theory calibrations.

Figure 1.

Change in PROMIS Fatigue T-scores Across Clinical Samples

CHF=Chronic heart failure; COPD=Chronic Obstructive Pulmonary Disease; Scores reported are on the T-score metric as referenced to the US general population (mean=50; SD=10).[34] Higher scores reflect more fatigue.

NOTE:

1. Cancer (group 1 n=81) : patients reported more (worse) fatigue at follow-up rated by the fatigue-specific global change item. Cancer (group 2 n=84) : patients reported less (better) fatigue at follow-up rated by the fatigue-specific global change item.

2. RA (group 1 n=171) :rhemautoid patients who reported more (worse) fatigue at follow-up rated by the fatigue-specific global change item. RA (group 2 n=48) :rhemautoid patients reported less (better) fatigue at follow-up rated by the fatigue-specific global change item.

Table 1

a. Mean Fatigue scores (with 95% CI) at Time 1 and Time 2 from mixed models

	Back Pain	Cancer	CHF	COPD Exacerbation	COPDStable	MDD	RA
Sample n (baseline; follow-up)	218; 173	310; 277	60;48	46; 44	79;72	196; 187	521;451
Baseline	56.7 (55.4, 58.0)	52.0 (51.1, 52.8)	58.8 (56.2, 61.4)	62.9 (60.4, 65.4)	56.1 (54.0, 58.2)	61.3 (60.0, 62.6)	53.8 (53.0, 54.5)
Follow-up	53.2 (51.8, 54.6)	53.1 (52.2, 54.0)	47.3 (44.5, 50.1)	56.3 (53.8, 58.8)	53.8 (51.6, 55.9)	55.6 (54.3, 56.9)	54.1 (53.4, 54.9)
Mean change	−3.47 (−4.7, −2.3)	1.16 (0.46, 1.85)	−11.56 (−14.4, −8.7)	−6.60 (−9.2, −4.0)	−2.36 (−3.7, −1.1)	−5.68 (−7.0, −4.4)	0.40 (−0.16, 0.95)

b. Mean Fatigue scores (with 95% CI) at Time 1 and Time 2 from mixed models – Cancer and RA groups

	Cancer – Improved	Cancer- about the same	Cancer - Worsened	RA – Improved	RA – abount the same	RA - Worsened
Sample n	84	81	81	48	232	171
Baseline	51.5 (50.1, 52.9)	50.5 (48.8, 52.2)	53.7 (52.0, 55.4)	51.8 (49.6, 54.0)	50.6 (49.6, 51.6)	57.4 (56.3, 58.5)
Follow-up	50.3 (48.9, 51.7)	50.5 (48.8, 52.2)	58.4 (56.7, 60.1)	51.5 (49.4, 53.7)	50.6 (49.6, 51.6)	58.9 (57.8, 59.9)
Mean change	−1.17 (−2.47, 0.12)	0 (−0.96, 0.96)	4.67 (3.33, 6.02)	−0.27 (−2.03, 1.49)	0 (−0.8, 0.8)	1.44 (0.56, 2.33)

Follow-up: 3 month assessment for all studies, except CHF (10 weeks) and RA (6 months)

NOTE: Patients were grouped into "improved", "about the same" and "worsened" based on the fatigue-specific global change item that they responded at follow-up.

CHF = Chronic Heart Failure

COPD = Chronic Obstructive Pulmonary Disease

MDD = Major Depressive Disorder

RA = Rheumatoid Arthritis

Table 2

Responsiveness to Self-reported Change Overall Health (General Global Change) or Fatigue (Fatigue-Specific Global Change)

	Back Pain		CHF		COPD: Exacerbation		COPD: Stable		MDD		Cancer		RA	
	SRM (n)	Mean Change (SD)	SRM (n)	Mean Change (SD)	SRM (n)	Mean Change (SD)	SRM (n)	Mean Change (SD)	SRM (n)	Mean Change (SD)	SRM (n)	Mean Change (SD)	SRM (n)	Mean Change (SD)
General Global Change														
Better	−0.69 (51)	−4.89** (7.06)	−1.18 (44)	−11.60** (9.86)	−0.98 (7)	−8.74* (8.93)	−0.74 (12)	−5.24* (7.09)	−0.69 (43)	−7.12** (10.29)	−0.22 (51)	−1.35 (6.26)	−0.50 (61)	−2.79** (5.55)
About the same	−0.40 (95)	−3.46** (8.61)	Na	Na	−0.97 (14)	−4.84* (4.98)	−0.40 (40)	−1.77* (4.47)	−0.69 (113)	−6.14** (8.86)	0.17 (132)	0.93 (5.51)	0.09 (297)	0.52 (5.93)
Worse	−0.04 (24)	−0.33 (7.63)	−0.30 (2)	−4.00 (13.24)	−0.73 (13)	−4.69* (6.40)	−0.20 (14)	−1.21 (5.91)	−0.42 (30)	−2.74* (7.79)	0.55 (84)	3.24** (5.87)	0.43 (92)	2.59** (6.09)
Fatigue-Specific Global Change														
Better			−1.23 (41)	−11.90** (9.65)	−1.19 (16)	−11.62** (9.80)	−0.94 (15)	−6.02* (6.40)			−0.20 (84)	−1.17 (5.97)	−0.06 (48)	−0.35 (6.05)
About the same			−0.39 (2)	−7.79 (19.94)	−0.79 (25)	−4.29** (5.44)	−0.30 (45)	−1.41* (4.70)			0.00 (81)	0.00 (4.32)	−0.01 (232)	−0.06 (6.18)
Worse			−0.51 (3)	−4.89 (9.49)	−0.10 (2)	−1.04 (10.46)	0.22 (8)	0.57 (2.58)			0.77 (80)	4.65** (6.03)	0.25 (171)	1.47* (5.88)

CHF = Chronic Heart Failure

COPD = Chronic Obstructive Pulmonary Disease

MDD = Major Depressive Disorder

RA = Rheumatoid Arthritis

Negative changes represents improvement in fatigue. Positive changes represent worsening fatigue.

SRM: standardized response mean. The standardized response mean is the ratio of the mean change to the standard deviation of that change [33].

*
  $p < 0.05$

**
  $p < 0.001$