



Published in final edited form as:

*J Clin Epidemiol.* 2016 May ; 73: 119–127. doi:10.1016/j.jclinepi.2015.08.036.

## Clinical Validity of PROMIS® Depression, Anxiety, and Anger across Diverse Clinical Samples

Benjamin D. Schalet<sup>1</sup>, Paul A. Pilkonis<sup>2</sup>, Lan Yu<sup>3</sup>, Nathan Dodds<sup>2</sup>, Kelly L. Johnston<sup>2</sup>, Susan Yount<sup>1</sup>, William Riley<sup>4</sup>, and David Cella<sup>1</sup>

<sup>1</sup>Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

<sup>2</sup>Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

<sup>3</sup>Departments of Medicine and Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

<sup>4</sup>National Cancer Institute (NCI), Bethesda, MD, USA

### Abstract

**Objective**—To evaluate the responsiveness to change of the PROMIS® negative affect measures (Depression, Anxiety, and Anger) using longitudinal data collected in six chronic health conditions.

**Study Design and Setting**—Individuals with major depressive disorder (MDD), back pain, chronic obstructive pulmonary disease (COPD), chronic heart failure (CHF), and cancer completed PROMIS negative affect instruments as computerized adaptive test (CAT) or as fixed-length short form (SF) at baseline and a clinically-relevant follow-up interval. Participants also completed global ratings of health. Linear mixed effects models and standardized response means (SRM) were estimated at baseline and follow-up.

**Results**—903 individuals participated (back pain,  $n = 218$ ; cancer,  $n = 304$ ; CHF,  $n = 60$ ; COPD,  $n = 125$ ; MDD,  $n = 196$ ). All three negative affect instruments improved significantly for treatments of depression and pain. Depression improved for CHF patients (anxiety and anger not administered), while anxiety improved significantly in COPD groups (stable and exacerbation).

---

Address correspondence to: Benjamin D. Schalet, Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Northwestern University, 633 N. St. Clair, 19th Floor, Chicago, IL 60611, Phone: 312-503-3640, Fax: 312-503-4800, b-schalet@northwestern.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

### CONFLICT OF INTEREST

Benjamin D. Schalet: None

Paul A. Pilkonis: None

Lan Yu: None

Nathan Dodds: None

Kelly L. Johnston: None

Susan Yount: None

William Riley: None

David Cella is an unpaid member of the board of directors and officer of the PROMIS Health Organization

Response to treatment was not assessed in cancer. Subgroups of patients reporting better or worse health showed a corresponding positive or negative average SRM for negative affect across samples.

**Conclusion**—This study provides evidence that the PROMIS negative affect scores are sensitive to change in intervention studies in which negative affect is expected to change. These results inform the estimation of meaningful change and enable comparative effectiveness research.

### Keywords

PROMIS; Depression; Anxiety; Anger; Chronic Conditions; Item Bank

---

## Introduction

Researchers and clinicians wishing to assess negative affect in a clinical or community population must choose from among numerous assessment options, many of which purport to measure the same or a similar construct[1–3]. Not all the available instruments meet high levels of instrument development standards for reliability, validity, appropriate reading level, and minimal respondent burden[4, 5]. In an effort to improve the existing measures, the Patient-Reported Outcomes Measurement Information System (PROMIS<sup>®</sup>) employed a multi-step, mixed methods approach to develop computerized adaptive tests (CAT) and fixed-length short forms to assess health-related quality of life, including symptoms and functional domains across physical, mental and social health [6]. Moreover, the goal of PROMIS, as an NIH Roadmap initiative, was to create a system that could standardize the measurement of patient-reported outcome across chronic conditions; thus, enabling comparisons of the burden of disease and the benefits of treatment across these chronic diseases. Included in that system is a set of item banks and short forms for negative affect, specifically depression, anxiety and anger[7].

This paper reports on an important subsequent step in the validation processes for PROMIS measures: longitudinal analysis of the PROMIS negative affect scores in adult samples of patients with specified chronic health conditions. These analyses have the potential to deepen the PROMIS validity base, help define anchor-based clinically-important differences, and further enable comparative effectiveness research by identifying subpopulation reference values and observed change scores based on receipt of conventional treatment. In the present study, these conditions comprised back pain, cancer, chronic obstructive pulmonary disease (COPD), chronic heart failure (CHF), and major depressive disorder (MDD). Self-reported negative affect may distinguish important features among patients suffering from these medical conditions, including level of risk, [8] disability, [9–11] or recovery. [12]

Although this investigation is an exploratory “test drive” of PROMIS measures, the nature of the clinical groups and interventions allows us to articulate some hypotheses. For each of the PROMIS negative affect measures, we hypothesized that longitudinal improvements would occur during treatment for MDD (psychotropic medications and/or psychotherapy), chronic heart failure (heart transplant surgery), back pain (spinal injection), and the resolution of COPD exacerbation. Further, we expected the greatest change on all three negative affect measures for those being treated for MDD relative to those being treated for physical

conditions. Given the progressive nature of cancer and the absence of any change in treatment for the COPD stable subgroup, we did not have a priori hypotheses for longitudinal changes of these groups. Our cross-sample hypotheses were that the MDD sample should have more severe scores on PROMIS Depression compared to those with other ailments, while patients with COPD exacerbation should have worse negative affect scores compared to the stable group.[9, 13, 14]

While we have articulated some hypotheses above, our ability to develop these more fully is somewhat hampered by the secondary nature of the data analysis. As discussed in the overview paper of this series, [15] a more thorough validation study developed with an a priori design, analytic approach, and data collection focused on across-study and across-disease validation would be useful and possibly more elegant. It should also be emphasized that the purpose of this report is not to demonstrate treatment effectiveness, but to investigate the responsiveness and validity of the PROMIS negative affect instruments.

## Method

### Measures

**PROMIS Depression, Anxiety, and Anger**—At the time of this study, there were three PROMIS negative affect item banks, consisting of Depression (28 items), Anxiety (29 items), and Anger (29 items). The items in the PROMIS negative affect banks use a 7-day time frame and a 5-point rating scale that ranges from 1 (“Never”) to 5 (“Always”) [6, 7]. Each item bank was developed using comprehensive mixed (qualitative and quantitative) methods [16, 17]. After confirming essential unidimensionality and fit to the graded response model,[18] items were calibrated with regard to their location (severity) and discrimination (ability to distinguish people at different levels of distress). This produced a bank of questions that can accurately measure levels of negative affect across its observed continuum, and provides the basis for innovative administration strategies such as CAT (in which item administration selection is based on responses to prior items) and short-forms targeted to the particular sample being assessed. Each item bank provided more information than conventional measures across a wider range of severity, ranging from normal to severely distressed [7].

The PROMIS Depression bank focuses on affective and cognitive manifestations of depression rather than somatic symptoms such as appetite, fatigue and sleep. PROMIS Anxiety content focuses on fear (e.g., worry, feelings of panic), anxious misery (e.g., dread), hyperarousal (e.g., tension, nervousness, restlessness), and somatic symptoms related to arousal (e.g., cardiovascular symptoms, dizziness). The Anger bank included items that were affective and cognitive, but also included indicators of behavioral activation and anger expression [7]. See <http://www.nihpromis.org/measures/domainframework1> for full definitions of these banks.

The administration format of the PROMIS measures differed slightly across the condition and disease groups evaluated in this project. For most studies, the banks were administered via CAT. For the cancer study, however, customized short-forms that predated the release of PROMIS short forms (Version 1.0) were administered.[10] We only scored items on the two

cancer short forms that are also in the PROMIS negative affect item banks [7]; eight items for depression and seven for anxiety. These items do differ, however, from the established short forms: for depression, six out of the eight items are also in the Depression Short Form 8b (SF-8b; [7]); for anxiety, two out of the seven items are also in the Anxiety Short Form 7a [7]. Nevertheless, because the cancer study items were taken from the PROMIS negative affect item banks, we could score them with the established PROMIS discrimination and threshold parameters; therefore, scores were estimated using the same T-score metric as the existing short forms and CATs. These cancer study short-forms correlated very highly with the full bank ( $r = .97$  for both anxiety and depression).

**Clinical Change Anchors**—For each data set, we identified one general health and one general emotional distress item (outside of PROMIS items) to serve as a clinical indicator of change. These items either assessed change over the time of treatment directly or via a calculated change score. For example, patients in the cancer study answered the following question at each administration, “In general, would you say your health is...” with five respondent choices ranging from “excellent” to “poor.” In this case, we subtracted the former from the later scores to determine self-reported change in general health. Cancer patients also answered a domain-specific question at the second administration, namely, “Since the last time you filled out a questionnaire, your level of anxiety is...” with seven respondent choices ranging from “very much better” to “very much worse.” Patient scores were categorized into three groups of patients, reflecting health changes that were better, about the same or worse. Details on the anchors used for each study are described in the overview paper in this issue. [15]

## Samples

The PROMIS negative affect measures were administered as part of the PROMIS studies designed to validate PROMIS measures in a variety of clinical populations. Five patient groups completed one or more forms of the PROMIS negative affect banks: 1) MDD, 2) back pain, 3) COPD, both exacerbating and stable patients, 4) CHF, and 5) cancer. The studies of MDD, back pain, and CHF followed patients as they enrolled in new treatments. COPD exacerbation patients were treated for their condition, which was expected to resolve over the course of the 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 transplantation (CHF), and 6–12 weeks after enrollment (cancer). (Although cancer and COPD-stable groups were not enrolled in new treatments, we apply the terms “baseline” and “follow-up” to all study groups for consistency.) The percentages for missing follow-up data for PROMIS instruments were as follows: 5% for MDD, 10% for cancer, and 20% for CHF, and 20% for Pain. For COPD, the overall missing data rate was 7% for PROMIS Depression and Anger and 8% for Anxiety. For COPD exacerbation, the percentage was 4% for all three measures; for COPD stable, the percentage was 9% for PROMIS Depression and Anger, and 10% for Anxiety. Detailed information on patient characteristics and treatments may be found in the overview article elsewhere in this journal. [15]

## Statistical Analyses

PROMIS Depression, Anxiety, and Anger were administered to 5, 4, and 3 different clinical groups, respectively. Least square means were estimated for these longitudinal data. Linear mixed models were estimated with random subject effects to account for the similarity among repeated observations within individuals [19, 20]. 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 those respondents with data at both time points [21, 22]. Least squares means, standard errors and 95% confidence intervals were estimated from the models.

To clinically anchor the changes in PROMIS negative affect measures, we used items that assessed changes in overall health, negative affect, or both, as described above. For each clinically anchored subgroup, we calculated the change in T-score and the standardized response mean (SRM). The latter is the ratio of the mean change to the standard deviation of that change, [23, 24] which is a form of Cohen's effect size index [25]. Based on previous studies, we assume for the purposes of this study that an SRM of .30 would be a minimally important difference in outcome [10, 26, 27]. Because of missing data at follow-up, subgroup sample sizes for the anchor-based analysis do not sum to the sample sizes for the mixed methods analysis.

Because the cancer sample was sufficiently large (enrolled  $N=310$ ) and some patients were expected to improve while others would deteriorate, we also computed least square means for the subsamples of patients who reported that their overall health got better, worse, or remained about the same over the course of the study.

## Results

### Longitudinal Change

The results of the mixed models are summarized in Table 1 and 2, while Figure 1 shows the least squares means of each measure in each clinical group. Consistent with our expectations, scores on the negative affect measures decreased significantly for patients receiving treatment for back pain, MDD, and CHF. The COPD exacerbation group, however, improved only on PROMIS anxiety and not (significantly) on anger and depression on the second administration following their exacerbation. Furthermore, the COPD stable group also showed significant reductions of anxiety (but not on Anger and Depression). As Figure 1 illustrates, both COPD groups were also more highly elevated on Anxiety at baseline relative to Anger and Depression.

In the cancer sample, we observed small improvements in mean levels of the entire group, but only the improvement on PROMIS Anxiety ( $-0.8$  T-score points) was significant. However, as Figure 1 and Table 2 show, the clinically anchored subgroups produced significant improvements in PROMIS Anxiety and Depression in the predicted direction. Those reporting global improvements on the domain-specific anchor ("Since the last time you filled out a questionnaire, your level of depression [anxiety] is...") also improved

significantly on PROMIS Anxiety and Depression; patients reporting deterioration on the domain-specific anchor also showed significantly higher Anxiety and Depression scores.

### Sample Differences

As expected, the mean score for PROMIS Depression (at baseline) was highest in the MDD sample compared to the other clinical samples ( 5 T-score points). The negative affect measures also distinguished the COPD stable group from COPD exacerbation for PROMIS Anxiety,  $t(119) = 3.68, p < .001$ , PROMIS Depression,  $t(119) = 2.76, p < .01$ , and marginally for PROMIS Anger,  $t(119) = 1.87, p = .06$ , with the COPD exacerbation group showing more distress. These group differences were maintained at follow-up (3 months) for Anxiety ( $t(110) = 2.85, p < .01$ ) and Depression ( $t(111) = 2.61, p < .05$ ), but the difference was not significant for Anger ( $t(111) = 1.06, p = .29$ ).

### Clinically Anchored Subgroups

The analysis of clinically anchored subgroups generally showed that individuals grouped in the better health category saw greater improvements on PROMIS negative affect compared to those in the worse health category (see Tables 3–5); this was particularly the case when the anchor was domain-specific (negative affect) compared to global (health). For PROMIS Depression, individuals reporting better health between administrations showed an average SRM across conditions of  $-0.54$  compared to an average of  $-0.10$  for those in the worse health group. Changes in the groups defined by the domain-specific distress anchor were even larger: the average SRM was  $-0.71$  for better compared to  $0.49$  for worse. For PROMIS Anxiety, the better health groups showed an average of  $-0.66$  across conditions compared to an average of  $-0.17$  for those reporting worse health. Using the distress anchor, the SRM averages were  $-0.83$  for better and  $0.38$  for worse. For Anger, the globally anchored groups showed less of a difference, with an SRM of  $-0.44$  for better health compared to  $-0.01$  for worse health. With the distress anchor, however, the average was  $-0.62$  for better health and  $0.56$  for worse health.

### Discussion

The PROMIS negative affect measures were created using an extensive instrument development process, including qualitative and cognitive interviews, [16] dimensionality analyses, IRT calibration, and concurrent validity evaluation, [6, 7] and scale-setting to match the US population[28]. The current study examines how well the PROMIS emotional distress instruments differentiate among diverse clinical groups and if changes over time are consistent with expected changes during treatment.

Our results were largely consistent with a priori hypotheses. As expected, we found statistically significant longitudinal reductions on negative affect measures after treatment for MDD (psychotropic medications and/or psychotherapy), chronic heart failure (heart transplantation), and back pain (spinal injection). In addition, we found the greatest change on all three emotional distress measures for the treatment of MDD relative to the physical conditions and other treatments. Our cross-sample hypotheses were also met: patients with MDD scored higher on the emotional distress measures compared to the other groups, while

COPD exacerbation patients scored higher on Depression and Anxiety than the stable group (but the two groups were not significantly different for Anger). We did not have a priori expectations for changes in cancer and COPD-stable groups; significant improvements were found in both groups on Anxiety.

While we expected that all three emotional distress measures would show improvement upon the resolution of COPD exacerbation, only PROMIS Anxiety improved for this group ( $-4.3$  T-score points). The COPD-stable group likewise improved significantly on PROMIS Anxiety ( $-3.0$  T-score points). For each COPD group, the PROMIS Anxiety score was also the highest among the three emotional distress measures at baseline. In fact, the mean level of anxiety at baseline in the COPD exacerbation group (60.2) was nearly as high as the level of anxiety in the MDD sample (61.7). Anxiety also clearly distinguished COPD exacerbation from COPD stable, both at baseline and follow-up (difference  $5$  T-score points). Our results are consistent with research suggesting a possible unique role for anxiety in COPD progression and treatment. The prevalence and severity of negative affect, especially anxiety, in people with COPD is well documented, [9, 13, 14] and anxiety in women with COPD is independently associated with increased risk of death. [29]

Mean scores on the emotional distress measures in the cancer group did not change much over the course of treatment ( $< 1$  T-score point). However, as this group consisted of a diverse set of participants in different stages and types of cancer (see overview paper elsewhere in this issue)[15] we would expect some individuals to improve and some to deteriorate. As Figure 1 illustrates, groups anchored to the global, domain-specific rating of improvement also showed significant parallel improvement on PROMIS Depression and Anxiety (and vice versa for the global, domain-specific rating of deterioration). This pattern of scores has been observed previously,[30] and appeared to be robust in these data, with the absolute value of SRMs ranging from .35 to .72 (Tables 3 and 4).

The analysis of SRMs by clinically anchored subgroups supported the hypothesis that PROMIS emotional health measures are responsive to change in diverse clinical groups. For the domain-specific ratings of change, all SRMs for the better and worse subgroups were higher than the MID value of .30 across the three emotional distress measures. The SRMs for subgroups determined by general health ratings, however, were somewhat inconsistent. While the better health subgroups showed improvement in the predicted direction on emotional distress, the worse health subgroups also showed improvement on the emotional distress measures (albeit below the .30 MID threshold in most cases). This discrepancy is plausible, as participants' general health may be influenced by factors external to treatment or otherwise independent from their emotional state.

Our results are also relevant to psychopathology research. Most research on the structure of the depressive and anxiety symptoms is done on cross-sectional samples; [31–33] however, longitudinal and treatment data can also support or refute models developed on cross-sectional data. Figure 1 illustrates that the course of depression and anxiety are virtually identical for the outpatients with MDD. This may reflect the correlation/comorbidity of depression and anxiety and suggests that it would be useful to investigate how they can be modeled and examined together [34].

The present study is not without limitations. First, while the use of global ratings of health or domain-specific change may be face valid and clinically relevant, they provide some methodological complications. Because retrospective ratings are assessed at follow-up, they are typically correlated more with follow-up (current) scores rather than pre-test or change scores[35]. Secondly, sample sizes for COPD and CHF were modest; when anchored to different change groups, several of the subsamples for COPD and CHF were below 20. Consequently, results for these subgroups should be interpreted cautiously. Third, as stated in the introduction, this paper represents a secondary analysis of a range of studies with divergent goals. One of these goals (not addressed here) was to compare the responsiveness of PROMIS measures to established legacy measures; for depression, this question has been addressed in a separate manuscript. [36] For the purposes of this particular report, however, it would have been ideal to have the research design (e.g., choice of population and instruments) be exclusively informed by a priori hypotheses in the negative affect domain. Finally, we acknowledge that the anger bank was administered to fewer clinical patients with less extreme anger scores (relative to depression and anxiety). It would be important to test the responsiveness of this bank in populations that have known anger problems.

Despite these limitations, the current study extends previous validity research on PROMIS negative affect measures by examining longitudinal change in a diverse set of clinical groups. The study demonstrates predictable change on PROMIS negative affect for various treated clinical samples over time, with parallel changes for clinically anchored subgroups, and meaningful differences between these clinical groups. These data help inform an evidence base for defining treatment responders and conducting or interpreting the results of comparative effectiveness research.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

PROMIS® was funded with cooperative agreements from the National Institutes of Health (NIH) Common Fund Initiative (Northwestern University, PI: David Cella, PhD, U54AR057951, U01AR052177, R01CA60068; Northwestern University, PI: Richard C. Gershon, PhD, U54AR057943; American Institutes for Research, PI: Susan (San) D. Keller, PhD, U54AR057926; State University of New York, Stony Brook, PIs: Joan E. Broderick, PhD and Arthur A. Stone, PhD, U01AR057948, U01AR052170; University of Washington, Seattle, PIs: Heidi M. Crane, MD, MPH, Paul K. Crane, MD, MPH, and Donald L. Patrick, PhD, U01AR057954; University of Washington, Seattle, PI: Dagmar Amtmann, PhD, U01AR052171; University of North Carolina, Chapel Hill, PI: Harry A. Guess, MD, PhD (deceased), Darren A. DeWalt, MD, MPH, U01AR052181; Children's Hospital of Philadelphia, PI: Christopher B. Forrest, MD, PhD, U01AR057956; Stanford University, PI: James F. Fries, MD, U01AR052158; Boston University, PIs: Alan Jette, PT, PhD, Stephen M. Haley, PhD (deceased), and David Scott Tulskey, PhD (University of Michigan, Ann Arbor), U01AR057929; University of California, Los Angeles, PIs: Dinesh Khanna, MD (University of Michigan, Ann Arbor) and Brennan Spiegel, MD, MSHS, U01AR057936; University of Pittsburgh, PI: Paul A. Pilkonis, PhD, U01AR052155; Georgetown University, PIs: Carol M. Moinpour, PhD (Fred Hutchinson Cancer Research Center, Seattle) and Arnold L. Potosky, PhD, U01AR057971; Children's Hospital Medical Center, Cincinnati, PI: Esi M. Morgan DeWitt, MD, MSCE, U01AR057940; University of Maryland, Baltimore, PI: Lisa M. Shulman, MD, U01AR057967; and Duke University, PI: Kevin P. Weinfurt, PhD, U01AR052186). NIH Science Officers on this project have included Deborah Ader, PhD, Vanessa Ameen, MD (deceased), Susan Czajkowski, PhD, Basil Eldadah, MD, PhD, Lawrence Fine, MD, DrPH, Lawrence Fox, MD, PhD, Lynne Haverkos, MD, MPH, Thomas Hilton, PhD, Laura Lee Johnson, PhD, Michael Kozak, PhD, Peter Lyster, PhD, Donald Mattison, MD, Claudia Moy, PhD, Louis Quatrano, PhD, Bryce Reeve, PhD, William Riley, PhD, Peter Scheidt, MD, Ashley Wilder Smith, PhD, MPH, Susana Serrate-Sztejn, MD, William Phillip Tonkins, DrPH, Ellen Werner, PhD, Tisha Wiley, PhD, and James Witter, MD, PhD. The contents of this article uses



data developed under PROMIS. These contents do not necessarily represent an endorsement by the US Federal Government or PROMIS. See [www.nihpromis.org](http://www.nihpromis.org) for additional information on the PROMIS® initiative.

## References

1. Nezu, AM.; Ronan, GF.; Meadows, EA., et al. Practitioner's Guide to Empirically Based Measures of Depression. Boston, MA: Springer US; 2002.
2. Roemer, L. Measures for anxiety and related constructs. In: Antony, MM.; Orsillo, SM.; Roemer, L., editors. Practitioner's Guide to Empirically-Based Measures of Anxiety. New York: Springer; 2002. p. 49-83.
3. Harrington, JL.; Antony, MM. Oxford handbook of anxiety and related disorders. Oxford; New York: Oxford University Press; 2008.
4. Vodermaier A, Linden W, Siu C. Screening for Emotional Distress in Cancer Patients: A Systematic Review of Assessment Instruments. *J Natl Cancer Inst.* 2009; 101(21):1464–1488. [PubMed: 19826136]
5. McHugh RK, Rasmussen JL, Otto MW. Comprehension of self-report evidence-based measures of anxiety. *Depress Anxiety.* 2011; 28(7):607–614. [PubMed: 21618668]
6. Cella D, Riley W, Stone A, 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]
7. Pilkonis PA, Choi SW, Reise SP, et al. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS):depression, anxiety, and anger. Assessment. 2011; 18(3):263–283. Epub June 21, 2011. [PubMed: 21697139]
8. Kubzansky LD, Cole SR, Kawachi I, et al. Shared and unique contributions of anger, anxiety, and depression to coronary heart disease: a prospective study in the normative aging study. *Ann Behav Med.* 2006; 31(1):21–29. [PubMed: 16472035]
9. Blinderman CD, Homel P, Andrew Billings J, et al. Symptom Distress and Quality of Life in Patients with Advanced Chronic Obstructive Pulmonary Disease. *J Pain Symptom Manage.* 2009; 38(1):115. [PubMed: 19232893]
10. Yost KJ, Eton DT, Garcia SF, et al. 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]
11. Rothrock N, Hays R, Spritzer K, et al. Relative to the general US population, chronic diseases are associated with poorer health-related quality of life as measured by the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Clin Epidemiol.* 2010; 63(11):1195–1204. [PubMed: 20688471]
12. Rundell SD, Sherman KJ, Heagerty PJ, et al. The Clinical Course of Pain and Function in Older Adults with a New Primary Care Visit for Back Pain. *J Am Geriatr Soc.* 2015; 63(3):524–530. [PubMed: 25754841]
13. Walke LM, Byers AL, Tinetti ME, et al. Range and severity of symptoms over time among older adults with chronic obstructive pulmonary disease and heart failure. *Arch Intern Med.* 2007; 167(22):2503–2508. [PubMed: 18071174]
14. Walke LM, Gallo WT, Tinetti ME, et al. The burden of symptoms among community-dwelling older persons with advanced chronic disease. *Arch Intern Med.* 2004; 164(21):2321–2324. [PubMed: 15557410]
15. Cook KF, Jensen SE, Schalet BD, 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.
16. DeWalt DA, Rothrock N, Yount S, et al. Evaluation of Item Candidates: The PROMIS Qualitative Item Review. *Medical Care.* 2007; 45(5 Suppl 1):S12–S21. [PubMed: 17443114]
17. Kelly MA, Morse JQ, Stover A, et al. Describing depression: congruence between patient experiences and clinical assessments. *Br J Clin Psychol.* 2011; 50(1):46–66. [PubMed: 21332520]

18. Samejima, F. Estimation of latent ability using a response pattern of graded scores. Richmond, VA: Psychometric Society; 1969. Available from: <http://www.psychometrika.org/journal/online/MN17.pdf>
19. Verbeke, G.; Molenberghs, G. Linear Mixed Models for Longitudinal Data. New York, NY: Springer-Verlag; 2000.
20. Hedeker, DR.; Gibbons, RD. Longitudinal data analysis. Hoboken, N.J.: Wiley-Interscience; 2006.
21. Troxel AB, Fairclough DL, Curran D, et al. Statistical analysis of quality of life with missing data in cancer clinical trials. *Stat Med.* 1998; 17(5–7):653–666. [PubMed: 9549814]
22. Little, RJA.; Rubin, DB. Statistical Analysis with Missing Data. Hoboken, NJ: John Wiley & Sons, Inc.; 2002.
23. Fayers, PM.; Machin, D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. 2. Chichester, England; Hoboken, NJ: John Wiley & Sons; 2007.
24. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care.* 1990; 28(7):632–642. [PubMed: 2366602]
25. Cohen, J. Statistical power analysis for the behavioral sciences. 2nd. Hillsdale, N.J.: L. Erlbaum Associates; 1988.
26. Yost KJ, Cella D, Chawla A, et al. Minimally important differences were estimated for the Functional Assessment of Cancer Therapy-Colorectal (FACT-C) instrument using a combination of distribution- and anchor-based approaches. *J Clin Epidemiol.* 2005; 58(12):1241–1251. [PubMed: 16291468]
27. Revicki DA, Cella D, Hays RD, et al. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes.* 2006; 4:1–5. [PubMed: 16393335]
28. Liu H, Cella D, Gershon R, et al. Representativeness of the Patient-Reported Outcomes Measurement Information System Internet Panel. *Journal of Clinical Epidemiology.* 2010; 63(11):1169–1178. [PubMed: 20688473]
29. Divo M, Cote C, de Torres JP, et al. Comorbidities and Risk of Mortality in Patients with Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med.* 2012; 186(2):155–161. [PubMed: 22561964]
30. 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]
31. Krueger RF. The structure of common mental disorders. *Arch Gen Psychiatry.* 1999; 56(10):921–926. [PubMed: 10530634]
32. Krueger RF, Markon KE. Reinterpreting comorbidity: a model-based approach to understanding and classifying psychopathology. *Annu Rev Clin Psychol.* 2006; 2:111–133. [PubMed: 17716066]
33. Griffith JW, Zinbarg R, Craske M, et al. Neuroticism as a common dimension in the internalizing disorders. *Psychol Med.* 2010; 40(7):1125–1136. [PubMed: 19903363]
34. Watson D. Differentiating the mood and anxiety disorders: a quadripartite model. *Annu Rev Clin Psychol.* 2009; 5:221–247. [PubMed: 19327030]
35. Norman G, Stratford P, Regehr G. Bias in retrospective calculation of responsiveness to change. *Journal of Clinical Epidemiology.* 1997; 8:869–879. [PubMed: 9291871]
36. Pilkonis PA, Yu L, Dodds NE, et al. Validation of the depression item bank from the Patient-Reported Outcomes Measurement Information System (PROMIS) in a three-month observational study. *J Psychiatr Res.* 2014; 56:112–119. [PubMed: 24931848]

**What is new?****Key Findings**

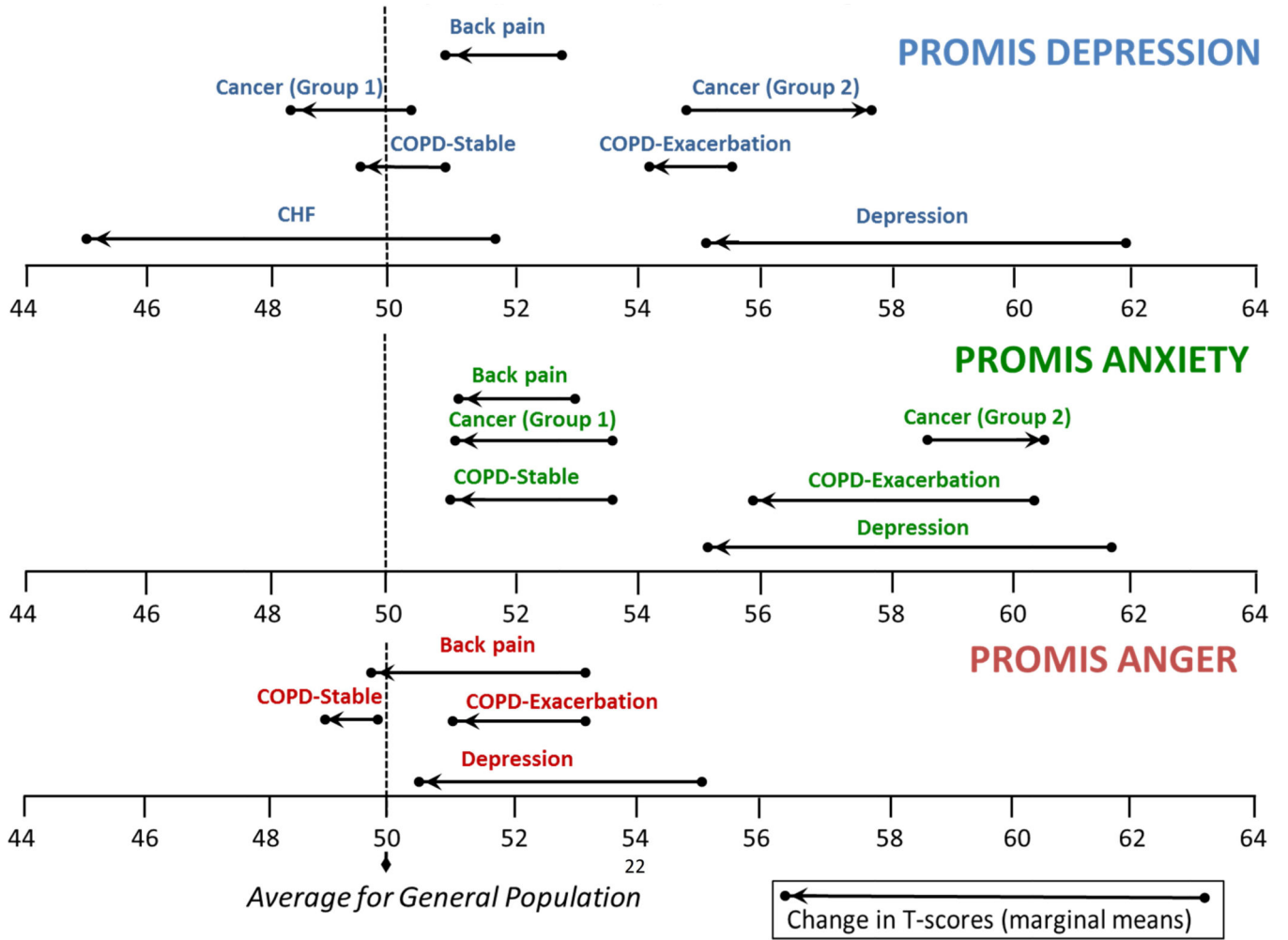
- PROMIS negative affect CATs and short-form instruments are predictably responsive in intervention studies that target depression, back pain and chronic heart failure

**What this adds to what was known?**

- A diverse set of clinical groups are differentiated by PROMIS negative affect scores

**What is the implication and what should change now?**

- With evidence of responsiveness and the ability to discriminate between clinical groups, PROMIS negative affect instruments are suitable for clinical trials and comparative effectiveness studies.



**Figure 1. PROMIS Negative Affect Scores in Different Clinical Groups over Time**  
 Higher scores indicate more negative affect. A T-score of 50 reflects the mean (and 10 the standard deviation) in the US general population sample that was used to center the T-scores. [28] Cancer (Group 1) patients reported improvement on depression/anxiety on the domain-specific global change rating. Cancer (Group 2) reported more depression/anxiety on domain-specific global change rating.  
*Abbreviations:* COPD, chronic obstructive pulmonary disease; CHF, chronic heart failure.

**Table 1**

PROMIS Negative Affect Scores in Different Clinical Samples

	Back Pain (n=218/174)	COPD: Stable (n=79/72)	COPD: Exacerbation (n=46/44 <sup>a</sup> )	MDD (n=196/187)	Cancer (n = 310/278)	CHF (n=60/48)
<b>Depression</b>						
<b>Baseline</b>	53.2 (52.0, 54.5)	51.0 (48.8, 53.1)	55.7 (52.6, 58.9)	61.9 (60.6, 63.2)	50.3 (49.4, 51.2)	51.7 (49.2, 54.2)
<b>Follow-up</b>	50.8 (49.5, 52.1)	49.4 (47.2, 51.6)	54.2 (51.0, 57.4)	55.2 (53.9, 56.5)	49.9 (49.0, 50.8)	44.9 (42.2, 47.7)
<b>Change</b>	-2.5 (-3.6, -1.4)**	-1.5 (-3.2, 0.1)	-1.5 (-4.4, 1.3)	-6.7 (-8.0, -5.5)**	-0.4 (-1.0, 0.3)	-6.8 (-10.0, -3.5)**
<b>Anxiety</b>						
<b>Baseline</b>	55.5 (54.3, 56.7)	53.8 (51.6, 56.1)	60.2 (57.3, 63.1)	61.7 (60.4, 62.9)	53.2 (52.3, 54.1)	--
<b>Follow-up</b>	52.8 (51.5, 54.1)	50.9 (48.6, 53.1)	55.9 (53.0, 58.9)	55.3 (54.1, 56.6)	52.4 (51.5, 53.3)	--
<b>Change</b>	-2.7 (-3.8, -1.6)**	-3.0 (-4.8, -1.1)*	-4.3 (-7.0, -1.5)*	-6.3 (-7.6, -5.0)**	-0.8 (-1.5, -0.1)*	--
<b>Anger</b>						
<b>Baseline</b>	53.3 (52.1, 54.4)	49.9 (47.6, 52.2)	53.2 (50.2, 56.3)	55.1 (53.7, 56.5)	--	--
<b>Follow-up</b>	49.7 (48.4, 50.9)	48.8 (46.5, 51.2)	50.9 (47.8, 54.0)	50.6 (49.1, 52.0)	--	--
<b>Change</b>	-3.6 (-4.8, -2.4)**	-1.0 (-2.9, 0.9)	-2.3 (-5.3, 0.7)	-4.5 (-5.9, -3.1)**	--	--

Entries in the table denote the least squares mean and 95% confidence interval, as estimated in the mixed models. Sample sizes represent the enrolled participants / follow-up.

Abbreviations: COPD, chronic obstructive pulmonary disease; CHF, chronic heart failure; MDD, major depressive disorder

<sup>a</sup> Follow-up sample size for COPD exacerbation was 43 for Anxiety.

\*  $p < 0.05$

\*\*  $p < 0.001$

**Table 2**

## PROMIS Depression/Anxiety Scores in Cancer Subsamples

	<b>Cancer-Better</b> ( <i>n</i> = 70)	<b>Cancer-Same</b> ( <i>n</i> = 135)	<b>Cancer-Worse</b> ( <i>n</i> = 38)
<b>Depression</b>			
<b>Baseline</b>	50.3 (48.6, 51.9)	49.6 (48.3, 50.9)	54.8 (52.9, 56.7)
<b>Follow-up</b>	48.2 (46.5, 49.9)	48.9 (47.6, 50.2)	57.8 (55.9, 59.7)
<b>Change</b>	-2.1 (-3.4, -0.7) *	-0.7 (-1.6, 0.2)	3.0 (1.7, 4.4) **
<b>Anxiety</b>			
	<i>n</i> = 76	<i>n</i> = 129	<i>n</i> = 41
<b>Baseline</b>	53.8 (52.2, 55.4)	51.7 (50.4, 53.0)	58.5 (56.6, 60.4)
<b>Follow-up</b>	51.1 (49.5, 52.7)	50.9 (49.7, 52.2)	60.4 (58.5, 62.3)
<b>Change</b>	-2.7 (-4.1, -1.2) **	-0.7 (-1.7, 0.2)	1.9 (0.2, 3.6) *

Entries in the table denote the least squares mean and 95% confidence interval, as estimated in the mixed models. Cancer subgroups were created using responses to the question, “Since the last time you filled out a questionnaire, your level of depression [anxiety] is...,” assessed at follow-up. The sum of the sample sizes for the change groups (*n* = 278) represents 78% of the enrolled cancer sample.

\*  
*p* < 0.05

\*\*  
*p* 0.001

**Table 3**  
Change in PROMIS Depression by General Health and Domain-Specific Anchors

Clinical Subgroup	PROMIS Depression Change by General Health Anchor			PROMIS Depression Change by Depression / Distress Anchor		
	n	SRM	T-Score (SD)	n	SRM	T-Score (SD)
<b>Cancer</b>						
Better	51	-0.26	-1.3 (4.9)	70	-0.37 <sup>a</sup>	-2.1 (5.6)
About the Same	131	-0.17	-1.0 (5.6)	134	-0.13	-0.7 (5.4)
Worse	85	0.13	0.7 (5.3)	38	0.72 <sup>a</sup>	3.0 (4.2)
<b>Chronic Heart Failure</b>						
Better	44	-0.61 <sup>a</sup>	-6.9 (11.3)	32	-0.71 <sup>a</sup>	-8.6 (12.1)
About the Same	0	--	--	12	-0.34 <sup>a</sup>	-2.6 (7.5)
Worse	2	-0.23	-2.3 (10.1)	1	--	4.8 --
<b>COPD - Exacerbation</b>						
Better	7	-0.30 <sup>a</sup>	-2.6 (8.6)	14	-0.94 <sup>a</sup>	-7.2 (7.6)
About the Same	14	0.18	2.1 (11.6)	22	-0.05	-0.4 (7.0)
Worse	13	-0.29	-1.8 (6.4)	7	0.36 <sup>a</sup>	4.8 (13.4)
<b>COPD - Stable</b>						
Better	13	-0.53 <sup>a</sup>	-4.1 (7.9)	18	-0.53 <sup>a</sup>	-4.1 (7.6)
About the Same	40	-0.29	-1.9 (6.3)	38	-0.38 <sup>a</sup>	-2.3 (6.0)
Worse	14	0.14	1.1 (8.1)	14	0.55 <sup>a</sup>	3.7 (6.7)
<b>Major Depressive Disorder</b>						
Better	43	-1.08 <sup>a</sup>	-9.7 (9.0)	155	-0.99 <sup>a</sup>	-8.5 (8.5)
About the Same	113	-0.80 <sup>a</sup>	-6.7 (8.4)	17	0.14	0.8 (5.5)
Worse	30	-0.28	-2.3 (8.3)	15	0.32 <sup>a</sup>	1.8 (5.6)
<b>Back Pain</b>						
Better	51	-0.46 <sup>a</sup>	-3.4 (7.4)	--	--	--
About the Same	95	-0.32 <sup>a</sup>	-2.4 (7.5)	--	--	--

Clinical Subgroup	PROMIS Depression Change by General Health Anchor			PROMIS Depression Change by Depression / Distress Anchor		
	<i>n</i>	<i>SRM</i>	T-Score (SD)	<i>n</i>	<i>SRM</i>	T-Score (SD)
Worse	24	-0.10	-0.7 (7.0)	--	--	--

For each clinical sample, the sum of the three change groups is smaller than the enrolled sample, due to missing instrument and anchor data at follow-up. We list here the percentage of the total change group relative to the enrolled sample, with the general health anchor followed by domain-specific health groups: COPD-Exacerbation (74%, 93%), COPD-Stable (85%, 89%), Back Pain (78%, NA), Depression (95%, 95%), CHF (77%, 75%), and Cancer (86%, 78%).

<sup>a</sup>SRM 0.30

Abbreviations: COPD, chronic obstructive pulmonary disease



**Table 4**

Change in PROMIS Anxiety by General Health and Domain-Specific Anchors

Clinical Subgroup	PROMIS Anxiety Change by General Health Anchor				PROMIS Anxiety Change by Anxiety / Distress Anchor			
	<i>n</i>	<i>SRM</i>	T-Score	(SD)	<i>n</i>	<i>SRM</i>	T-Score	(SD)
<b>Cancer</b>								
Better	51	-0.22	-1.2	(5.7)	75	-0.43 <sup>a</sup>	-2.7	(6.3)
About the Same	131	-0.25	-1.5	(5.8)	129	-0.13	-0.7	(5.6)
Worse	85	0.07	0.4	(6.1)	41	0.35 <sup>a</sup>	1.9	(5.4)
<b>COPD - Exacerbation</b>								
Better	7	-1.19 <sup>a</sup>	-8.4	(7.1)	14	-1.28 <sup>a</sup>	-8.4	(6.5)
About the Same	14	-0.24	-2.4	(9.7)	22	-0.45 <sup>a</sup>	-4.0	(8.9)
Worse	13	-0.62 <sup>a</sup>	-3.2	(5.1)	7	0.25	2.3	(9.5)
<b>COPD - Stable</b>								
Better	13	-0.51 <sup>a</sup>	-4.2	(8.1)	18	-0.78 <sup>a</sup>	-5.7	(7.3)
About the Same	39	-0.53 <sup>a</sup>	-3.9	(7.3)	37	-0.46 <sup>a</sup>	-3.7	(8.1)
Worse	14	-0.02	-0.2	(9.4)	14	0.54 <sup>a</sup>	2.7	(5.0)
<b>Major Depressive Disorder</b>								
Better	43	-0.89 <sup>a</sup>	-9.6	(10.8)	--	--	--	--
About the Same	113	-0.76 <sup>a</sup>	-6.5	(8.5)	--	--	--	--
Worse	30	-0.19	-1.3	(6.9)	--	--	--	--
<b>Back Pain</b>								
Better	51	-0.49 <sup>a</sup>	-3.9	(8.0)	--	--	--	--
About the Same	95	-0.36 <sup>a</sup>	-2.5	(6.9)	--	--	--	--
Worse	24	-0.11	-0.8	(7.9)	--	--	--	--

For each clinical sample, the sum of the three change groups is smaller than the enrolled sample, due to missing instrument and anchor data at follow-up. We list here the percentage of the total change group relative to the enrolled sample, with the general health anchor followed by domain-specific health groups: COPD –Exacerbation (74%, 93%), COPD-Stable (84%, 87%), Back Pain (78%, NA), Depression (95%, NA), and Cancer (86%, 79%).

<sup>a</sup>SRM 0.30

Abbreviations: COPD, chronic obstructive pulmonary disease

**Table 5**  
Change in PROMIS Anger by General Health and Domain-Specific Anchors

Clinical Subgroup	PROMIS Anger Change by General Health Anchor			PROMIS Anger Change by Anger / Distress Anchor		
	n	SRM	T-Score (SD)	n	SRM	T-Score (SD)
<b>COPD - Exacerbation</b>						
Better	7	-0.49 <sup>a</sup>	-4.9 (10.1)	14	-0.78 <sup>a</sup>	-6.7 (8.5)
About the Same	14	-0.14	-1.7 (12.9)	22	-0.29	-2.4 (8.2)
Worse	13	0.07	0.5 (7.6)	7	0.50 <sup>a</sup>	6.5 (12.9)
<b>COPD - Stable</b>						
Better	13	-0.17	-1.7 (9.9)	18	-0.46 <sup>a</sup>	-3.8 (8.2)
About the Same	40	-0.23	-1.8 (7.7)	38	-0.25	-1.9 (7.7)
Worse	14	0.13	0.9 (6.9)	14	0.61 <sup>a</sup>	4.1 (6.6)
<b>Depression</b>						
Better	43	-0.64 <sup>a</sup>	-6.0 (9.3)	--	--	--
About the Same	113	-0.63 <sup>a</sup>	-5.5 (8.8)	--	--	--
Worse	30	0.08	1.0 (11.7)	--	--	--
<b>Back Pain</b>						
Better	51	-0.46 <sup>a</sup>	-3.4 (7.4)	--	--	--
About the Same	95	-0.52 <sup>a</sup>	-4.3 (8.3)	--	--	--
Worse	24	-0.32 <sup>a</sup>	-2.4 (7.5)	--	--	--

For each clinical sample, the sum of the three change groups is smaller than the enrolled sample, due to missing instrument and anchor data at follow-up. We list here the percentage of the total change group relative to the enrolled sample, with the general health anchor followed by domain-specific health groups: COPD - Exacerbation (74%, 93%), COPD - Stable (85%, 87%), Back Pain (78%, NA), Depression (95%, NA).

<sup>a</sup>SRM 0.30

Abbreviations: COPD, chronic obstructive pulmonary disease