Title

Responsiveness of 8 Patient-Reported Outcomes Measurement Information System (PROMIS) measures in a large, community-based cancer study cohort

Journal

Cancer, 123(2)

ISSN

0008-543X

Authors

Jensen, Roxanne E
Moinpour, Carol M
Potosky, Arnold L
et al.

Publication Date

2017-01-15

DOI

10.1002/cncr.30354

Peer reviewed

Original Article

Responsiveness of 8 Patient-Reported Outcomes Measurement Information System (PROMIS) Measures in a Large, Community-Based Cancer Study Cohort

Roxanne E. Jensen, PhD[1,2]; Carol M. Moinpour, PhD[3]; Arnold L. Potosky, PhD[1,2]; Tania Lobo, MS[2]; Elizabeth A. Hahn, MA[4]; Ron D. Hays, PhD[5]; David Cella, PhD[4]; Ashley Wilder Smith, PhD, MPH[6]; Xiao-Cheng Wu, MD, MPH[7]; Theresa H. M. Keegan, PhD, MS[8]; Lisa E. Paddock, PhD, MPH[9,10]; Antoinette M. Stroup, PhD[9]; and David T. Eton, PhD[11]

BACKGROUND: The Patient-Reported Outcomes Measurement Information System (PROMIS) was a National Institutes of Health-funded initiative to develop measures of symptoms and function. Responsiveness is the degree to which a measure can detect underlying changes over time. The objective of the current study was to document the responsiveness of 8 PROMIS measures in a large, population-based cancer cohort. **METHODS:** The Measuring Your Health study recruited 2968 patients who were diagnosed with 1 of 7 cancers between 2010 and 2012 through 4 Surveillance, Epidemiology, and End Results registries. Participants completed a baseline survey (6-13 months after diagnosis) and a 6-month follow-up survey. Changes in 8 PROMIS scores were compared with global ratings of transition, changes in performance status, and clinical events. **RESULTS:** Measures were responsive to 6-month declines and improvements in performance status with small to large effect sizes (ES) (Cohen d = 0.34-0.71; $P < .01$). Mean changes and effect sizes were larger for participants who reported declines compared with those who reported improvements. Small-to-medium ES were observed in patients who reported being "a little" worse (d = 0.31-0.56), and medium-to-large ES were observed in those who reported being "a lot" worse (d = 0.53-0.72). Hospitalized participants reported significant score increases, resulting in worsening of pain (d = 0.51), fatigue (d = 0.35), and depression (d = 0.57; all $P < .01$). Cancer recurrence and progression were associated with smaller increases in pain, fatigue, and sleep disturbance (d = 0.22-0.27). **CONCLUSIONS:** The current results indicated that all 8 PROMIS measures were sensitive to patient-perceived worsening and improvement and to major clinical events. These findings will be able to inform the design and interpretation of future research studies and clinical initiatives administering PROMIS measures. ***Cancer* 2016;000:000–000.** © *2016 American Cancer Society.*

KEYWORDS: oncology, patient-reported outcomes, responsiveness, validation studies..

INTRODUCTION

Patient-reported outcomes (PROs) are measures of functioning and well-being in physical, mental and social spheres of health.[1] In 2004, the National Institutes of Health (NIH) launched the Patient-Reported Outcomes Measurement Information System (PROMIS) as part of an NIH Roadmap (and, later, an NIH Common Fund) initiative assessing all disease and health areas.[2,3] PROMIS was designed using extensive qualitative and quantitative methods to develop a comprehensive set of item banks and short-form measures.

Responsiveness is an important aspect of scale evaluation for determining the degree to which a PRO measure can detect underlying true changes.[4] Anchors used as indicators of change include patient transition reports and documented

Corresponding author: Roxanne E. Jensen, PhD, Cancer Prevention and Control Program, Lombardi Comprehensive Cancer Center, Georgetown University, 3300 Whitehaven Street NW, Suite 4100, Washington, DC 20007; Fax: (202) 687-8444; rj222@georgetown.edu

[1]Department of Oncology, Georgetown University, Washington, DC; [2]Cancer Prevention and Control Program, Lombardi Comprehensive Cancer Center, Washington, District of Columbia; [3]Fred Hutchinson Cancer Research Center, Seattle, Washington; [4]Northwestern University Feinberg School of Medicine, Chicago, Illinois; [5]University of California-Los Angeles Department of Medicine, Los Angeles, California; [6]Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland; [7]Lousiana State University Health Sciences Center School of Public Health, Louisiana Tumor Registry, New Orleans, Louisiana; [8]Department of Internal Medicine, Division of Hematology and Oncology, University of California-Davis School of Medicine, Sacramento, California; [9]Cancer Epidemiology Services, New Jersey Department of Health, Trenton, New Jersey; [10]Rutgers School of Public Health and Cancer Institute of New Jersey, New Brunswick, New Jersey; [11]Division of Health Care Policy and Research, Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota

We thank the Measuring Your Health (MY-Health) Clinical Advisory Committee (Drs. Patricia Ganz, Julie Gralow, Maurie Markman, Jimmy Hwang, Anthony Back, and Bonnie Teschendorf); the Surveillance, Epidemiology, and End Results study site collaborators (Laura Allen; Lauren S. Maniscalco, MPH; Lisa Moy, MPH; Natalia Herman, MPH; Rosemary Cress, DrPH; and Wendy Ringer); and the Georgetown MY-Health Research team (Caroline Moore; Charlene Kuo, MPH; Marin Rieger, MS; Deena Loeffler, MA; and Aaron Roberts).

The Patient-Reported Outcomes Measurement Information System (PROMIS) is a National Institutes of Health Roadmap initiative to develop valid and reliable patient-reported outcome measures to be applicable across a wide range of chronic diseases and demographic characteristics. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

DOI: 10.1002/cncr.30354, **Received:** June 11, 2016; **Revised:** August 23, 2016; **Accepted:** August 24, 2016, **Published online** Month 00, 2016 in Wiley Online Library (wileyonlinelibrary.com)

clinical events over the study period. To date, only a few studies have been published examining the responsiveness of PROMIS measures.[5-7] One study examined PROMIS responsiveness and minimally important differences in a clinic-based sample of patients with advanced-stage cancer (n = 101).[8]

Our current study builds on these findings, evaluating the responsiveness of 8 PROMIS short-form measures in the Measuring Your Health (MY-Health) Study.[9] MY-Health was designed to conduct a large-scale psychometric evaluation of PROMIS measures across a diverse cancer sample[10] using community-based sampling to represent the full range of known health disparities across age and race/ethnic groups.[11] This study presents an ideal environment for evaluating responsiveness in a large sample of patients with 7 different cancers, providing 6-month prospective data and capturing both patient and clinical indicators. Demonstrating responsiveness in this heterogeneous cohort of patients with cancer will support using PROMIS measures in population-based studies, comparative effectiveness research, clinical trials, and other longitudinal studies.

MATERIALS AND METHODS

Study Sample and Data-Collection Procedures

We recruited patients with Cancer as part of the MY-Health study. Four population-based cancer registries, which are part of the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program in 3 states (California, Louisiana, and New Jersey), identified participants within 6 to 13 months after they were diagnosed with primary colorectal cancer, lung cancer, non-Hodgkin lymphoma, breast cancer, gynecologic (uterine, cervical) cancer, or prostate cancer. We oversampled by younger age (ages 21-49 and 50-64 years) and nonwhite race/ethnicity (Hispanic, black, and Asian), targeting 20% of the full sample represented (n = 1000) in each group. Enrolled participants completed a paper baseline survey and a 6-month follow-up survey. Among participants who completed our follow-up survey, we conducted a medical record abstraction (MRA) of cancer-related procedures, hospitalizations, and medical events for a random subsample. Further details on study design, eligibility, and baseline data-collection procedures have been described in depth elsewhere.[12]

Trained SEER study abstractors conducted an MRA on a 40% subsample of participants who completed both baseline and follow-up surveys. Patients with stage III and IV cancer were oversampled to ensure that this group had sufficient events for evaluation between the baseline and follow-up surveys. Abstractors reviewed hospital and outpatient records for cancer-related treatment (chemotherapy, hormone therapy, targeted therapies, radiation, and surgical procedures), hospitalizations, medical events, cancer status (recurrence, progression, remission), and vital status.

Clinical and Survey Variables

Data on the date of cancer diagnosis, cancer type, and cancer stage were obtained from routine SEER registry databases. Cancer stage was defined using American Joint Commission on Cancer criteria. Information on hospitalization (dates) and documented cancer status change (type and date of change: remission, recurrence, progression, missing) was collected through medical record review. The participant baseline survey collected the patient-reported demographic and clinical information (surgery, radiation, chemotherapy, hormone therapy) used in this study, including age, race/ethnicity, education level, and birthplace (United States or outside the United States). The race/ethnicity categories (non-Hispanic white [white], black, Hispanic, Asian) used for analysis were created following US Census (2010) classification algorithms.[13] When self-reported race/ethnicity was missing, we used information from the SEER registry database (<0.4% of participants).

Two non-PROMIS PRO measures were collected at both baseline and follow-up to examine PROMIS 6-month responsiveness: 1) the 7-item Functional Assessment of Cancer Treatment-General (FACT-G) Physical Well Being subscale (the full measure was not administered because of the overall survey length);[14] and 2) the single-item, patient self-report Eastern Cooperative Oncology Group performance status (ECOG PS) scale, which is often used in cancer clinical trials to assess disease impact on activities of daily living.[15] The ECOG PS scale has 5 response options ranging from "normal activity without symptoms" to "unable to get out of bed."

Main Outcome Measures

We created 8 PROMIS measures administered at baseline and follow-up time points: physical function (15 items), fatigue (14 items), pain interference (11 items), anxiety (11 items), depression (10 items), ability to participate in social roles version 2 (social function; 10 items), cognitive function version 2 (8 items), and sleep disturbance (8 items).[16,17] Each measure is a custom short form that was created for the MY-Health Study. These measures were designed to include multiple "off-the-shelf" PROMIS

short forms as well as frequent computer adaptive testing selections for lower functioning patients (0.5 and 1.0 standard deviations below the US population mean).

On the follow-up survey, a single-item patient-reported global rating of change was administered immediately after each of the 8 PROMIS measures.[18] This global change item asked the following: "Compared with 6 months ago, how is your (PROMIS symptom or function) now?" Two different response sets were used: 1) responses to pain, fatigue, and anxiety of "a lot less," "a little less," "about the same," "a little more," or "a lot more"; and 2) responses to physical function, social function, cognitive function, depression (labeled "feelings"), and sleep disturbance (labeled "sleep quality") of "a lot better," "a little better," "about the same," "a little worse," or "a lot worse."

Statistical Analyses

First, we evaluated our follow-up survey completion rate, examining demographic and clinical differences compared with the full baseline sample. Then, we conducted descriptive analyses of our sample to examine the 6-month change between our baseline and follow-up surveys for each PROMIS measure. We calculated means, standard deviations, and effect sizes for baseline and follow-up mean scores. Unadjusted change scores (baseline and 6-months scores) were evaluated within each anchor response option (self-reported measures of 6-month change) for each PROMIS measure.

We evaluated 6-month responsiveness across the 8 PROMIS measures using retrospective ratings of global change. Retrospective anchors were collected using self-reported 5-point global change ratings corresponding to each PROMIS domain measure. For each PROMIS measure, we calculated the mean score change and effect size across each change rating (eg, "a lot better").

We analyzed prospective change using 5 known-group contrasts based on both survey self-report (ECOG PS and current cancer status) and medical record information (the number of hospitalizations and recurrences/progressions vs remissions). Each comparison examined relative mean score differences between a group for which change was expected versus a stable contrast group for which no expected change was anticipated. We also evaluated Spearman rank-order correlations (r) of change for PROMIS measures (selected with baseline correlations above $r = 0.7$: Pain, Fatigue, Social Function, and Physical Function) using with the FACT Physical Well Being subscale.

We calculated the effect sizes of all responsiveness calculations by dividing the absolute value of the mean 6-month change in each PROMIS score by the baseline standard deviation, reported as the absolute value. We applied the Cohen interpretation of effect size for magnitude (Cohen d) as follows: $d = 0.2$ (small effect size), $d = 0.5$ (moderate effect size), $d = 0.8$ (large effect size).[19] Past studies suggest that change scores of approximately $d = 0.2$ are probably too low to be classified as an estimate of clinically meaningful change.[18,20,21] Therefore, we considered values at or above $d = 0.3$ as clinically meaningful change. We used the SAS version 9.4 statistical software package (SAS Institute Inc, Cary, NC).

RESULTS

The overall follow-up survey response rate was 54%. Participants who completed the follow-up survey were more likely to be non-Hispanic whites (62%), aged ≥ 65 years (57%), or to report a college degree or higher (62%). Patient with prostate cancer had the highest follow-up rates (61%), and those with cervical cancer had lowest (34%). The follow-up rate varied by cancer stage at diagnosis, decreasing as stage increased (57% stage I vs 46% stage IV). (Table 1). Overall, follow-up survey respondents reported significantly better function and lower symptom severity at baseline than those who were lost to follow-up.

Baseline characteristics of the participants who completed both a baseline and 6-month follow-up MY-Health survey ($n = 2968$) and had supplemental medical record information ($n = 844$) are provided in Table 1. Most patients were aged ≥ 50 years (81%), and 53% of respondents were members of a racial/ethnic minority group. Although most participants were diagnosed with breast and prostate cancer, 7 different cancer types were represented in the sample, and 780 patients (26%) were diagnosed with advanced disease. One-half of participants reported "normal" performance status at baseline. The medical record participant subset reflects an intentional oversampling of patients with advanced-stage cancer (stage III/IV).

Overall, we observed that participants were largely unchanged or had small improvements across all measures over a 6-month period. Among the PROMIS symptom measures, we observed small declines in pain interference (mean change, -1.6; $P < .001$) and fatigue (mean change, -1.1; $P < .001$). Improvements were observed for physical function and social function (mean change, 0.8 and 1.1, respectively; $P < .001$ for both). An analysis of self-reported, 6-month, retrospective global change indicated that more participants reported improvement (a little or a lot) for pain (51%), fatigue (49%), anxiety (47%), and social function (44%). However, symptom

TABLE 1. Patient Demographics

Variable	Baseline Sample		Follow-Up Cohort			Medical Record Subsample	
	No.	Column %	No.	Column %	Response Rate, %	No.	Column %
All	5506	100	2968	100	53.9	844	100
Age at diagnosis, y							
21-49	1203	21.8	572	19.3	47.5	153	18.1
50-64	2037	37	1111	37.4	54.5	357	42.3
65-84	2266	41.2	1285	43.3	56.7	334	39.6
Race/ethnicity							
White	2261	41.1	1394	47	61.7	392	46.4
Black	1121	20.4	600	20.2	53.5	164	19.4
Hispanic	1064	19.3	463	15.6	43.5	155	18.4
Asian	887	16.1	425	14.3	47.9	120	14.2
Other/multiple	173	3.1	86	2.9	49.7	13	1.5
Education attainment							
<HS graduate	981	17.8	438	14.8	44.6	143	16.9
HS graduate/Some College	2827	51.3	1499	50.5	53	417	49.4
College degree or higher	1622	29.5	1004	33.8	61.9	278	32.9
Missing/unknown	76	1.4	27	0.9	35.5	6	0.7
Born in United States							
Yes	3854	70	2211	74.5	57.4	626	74.2
No	1595	29	723	24.4	45.3	210	24.9
Missing	57	1	34	1.1	59.6	8	0.9
Cancer site							
Breast	1662	30.2	934	31.5	56.2	207	24.5
Cervix	149	2.7	51	1.7	34.2	9	1.1
Colorectal	937	17	493	16.6	52.6	192	22.7
Lung	722	13.1	309	10.4	42.8	113	13.4
NHL	464	8.4	261	8.8	56.3	101	12
Prostate	1177	21.4	718	24.2	61	170	20.1
Uterus	395	7.2	202	6.8	51.1	52	6.2
Stage							
I	1983	36	1127	38	56.8	190	22.5
II	1731	31.4	952	32.1	55	187	22.2
III	935	17	490	16.5	52.4	289	34.2
IV	635	11.5	290	9.8	45.7	176	20.9
Missing/unknown	222	4	109	3.7	49.1	2	0.2
ECOG PS at baseline							
Normal	2465	44.8	1488	50.1	60.4	376	44.5
Some symptoms	2054	37.3	1063	35.8	51.8	313	37.1
<50% Bed rest	681	12.4	313	10.5	46	120	14.2
>50% Bed rest	243	4.4	79	2.7	32.5	33	3.9
Missing	63	1.1	25	0.8	39.7	2	0.2
Cancer status: Medical record							
Remission			435	14.7		435	51.5
Never cancer free			143	4.8		143	16.9
Recurrence or progression			126	4.2		126	14.9
Missing/unknown			2264	76.3		140	16.6
Cancer status: Self-report at follow-up							
Cancer free			1649	55.6		418	49.5
Never cancer free			283	9.5		114	13.5
Recurrence or progression			77	2.6		24	2.8
Missing/unknown			959	32.3		288	34.1
Hospitalization: Medical record							
No			—	—		538	63.7
Yes			—	—		38	4.5
Missing			—	—		268	31.8

Abbreviations: ECOG PS, Eastern Cooperative Oncology Group performance status; NHL, non-Hodgkin lymphoma; HS. high school.

increases/functional declines reported over this period were highest for fatigue (13%) and lowest for physical function (7%) (Table 2).

Across symptom and functional status, an answer of "a lot better/less" was associated with mean changes of 2 to 4 points (d range = 0.22-0.44); "a little worse/more"

TABLE 2. Mean Symptom and Function Changes Over 6 Months and Distribution of Patient-Reported Change

Variable	Baseline Scores			Follow-Up Scores		Score Change		Retrospective (Patient-Reported) Change: "Past 6 Months," %				
	No.	Mean	SD	Mean	SD	Change	*P*	A Lot Less/Better	A Little Less/Better	About the Same	A Little More/Worse	A Lot More/Worse
Symptoms												
Pain interference	2965	51.7	10.6	50.1	10.6	−1.6	< .0001	39	12	32	7	4
Fatigue	2961	50.8	10.7	49.7	11.1	−1.1	< .0001	31	18	36	9	5
Depression	2946	47.5	10.3	47.7	10.7	0.2	.30	30	14	42	7	3
Anxiety	2946	48.5	10.7	48.7	11.2	0.2	.21	33	15	41	6	3
Sleep disturbance	2950	50.1	10.1	50.1	10.4	0.1	.63	20	16	52	8	3
Function												
Physical function	2953	46.1	9.7	46.9	10.1	0.8	< .0001	26	14	50	5	2
Social function	2949	51.5	10.6	52.6	10.8	1.1	< .0001	28	16	43	7	3
Cognitive function	2944	52.8	11.3	52.6	11.5	−0.2	.13	23	13	50	8	3

Abbreviation: SD, standard deviation.

was associated with mean changes of 3 to 6 points (d range = 0.31-0.56); and "a lot worse" was associated with mean changes of 5 to 9 points (d range = 0.53-0.72) (Table 3). Mean PROMIS changes and effect sizes were larger for respondents who reported declines in function or worsening symptoms. For example, the mean change from baseline in the PROMIS fatigue score was 5.42 (d = 0.62) among those who reported "a lot more fatigue" and −3.26 (d = 0.38) among those who reported "a lot less fatigue" (Table 3). Depression and cognitive function measures were the most responsive to declines (8.5 and 8.7, respectively) and the least responsive to improvements (−2.42 and 2.12, respectively) (Table 3).

Patients who had a 1-point improvement on the ECOG PS had clinically meaningful (d ≥ 0.30) improvements on 5 PROMIS measures. Physical function had the largest improvement (mean change, 3.4 points; d = 0.53), followed by pain (mean change, −4.5; d = 0.45) (Table 4). All measures were more sensitive to worsening performance status, with the largest change and effect size demonstrated for fatigue (mean change, 4.3; d = 0.63). Data from our MRA cohort indicated that pain, fatigue, and depression were responsive to hospitalization, with reported mean score increases of 3 to 5 points and small-to-moderate effect sizes (fatigue, d = 0.35; pain, 0.51; depression, 0.57). The correlation of change between the FACT-G Physical Well Being subscale ranged from $r = 0.33$ (pain) to $r = 0.47$ (fatigue) (Table 5).

DISCUSSION

The current study provides support for the responsiveness of 8 PROMIS measures in a diverse cohort of patients with cancer, supporting past evaluations in cancer and other chronic conditions.[22] Most notably, our study supports and extends research findings from a study of patients with advanced-stage cancer in which a similar range for longitudinal anchor-based change was identified in PROMIS pain, fatigue, anxiety, depression, and physical function measures (d = 0.36-0.67).[8] By using a larger, more diverse, national sample of patients with cancer, our findings provide evidence that a change of 3 to 5 points is sufficient across all PROMIS measures to identify clinically meaningful change.

This study also establishes that these PROMIS measures are responsive to functional recovery and symptom improvement in cancer. However, absolute changes in PROMIS scores tended to be smaller for patients who *retrospectively* reported a functional improvement/symptom decrease rather than for patients who reported a functional decline/symptom increase on global change ratings. This imbalance in change score magnitudes across retrospective

TABLE 3. Retrospective Anchor-Based Change Over 6 Months

Ratings Change Category by Measure	No.	Mean	SD	Baseline SD	Effect Size
Symptoms*					
Pain interference					
A lot less	1139	−3.74	9.31		0.38
A little less	356	−0.08	8.10		0.01
About the same	952	−1.15	8.12	10.50	0.11
A little more	210	3.74	9.55		0.37
A lot more	116	5.04	9.18		0.53
Fatigue					
A lot less	916	−3.26	7.69		0.38
A little less	516	−0.95	6.68		0.11
About the same	1074	−0.58	6.64	9.78	0.06
A little more	252	3.38	7.04		0.35
A lot more	134	5.42	8.03		0.62
Anxiety					
A lot less	884	−2.20	8.48		0.23
A little less	427	0.70	8.23		0.08
About the same	1247	0.29	7.61	10.69	0.03
A little more	213	5.02	7.81		0.48
A lot more	93	6.57	10.41		0.56
Depression					
A lot better	958	−2.42	8.02		0.27
A little better	428	1.14	8.21		0.13
About the same	1193	0.30	7.22	10.25	0.03
A little worse	170	5.61	8.05		0.56
A lot worse	75	8.70	9.21		0.72
Sleep disturbance					
A lot better	573	−1.97	6.08		0.29
A little better	458	−0.56	5.96		0.09
About the same	1520	0.36	5.63	8.12	0.05
A little worse	229	3.04	5.76		0.39
A lot worse	97	4.77	8.24		0.57
Function**					
Physical function					
A lot better	821	2.90	6.69		0.34
A little better	470	1.01	5.46		0.14
About the same	1274	0.42	5.61	9.73	0.04
A little worse	219	−3.02	5.38		0.37
A lot worse	80	−6.01	7.41		0.59
Social function					
A lot better	768	4.12	8.05		0.45
A little better	423	0.68	7.39		0.08
About the same	1464	0.45	7.59	10.53	0.04
A little worse	151	−3.20	8.91		0.31
A lot worse	70	−5.60	10.91		0.54
Cognitive function					
A lot better	693	2.12	8.21		0.22
A little better	384	0.05	8.24		0.02
About the same	1489	−0.21	8.10	11.24	0.02
A little worse	228	−4.99	8.88		0.45
A lot worse	81	−8.56	11.15		0.70

Abbreviations: SD, standard deviation.
*Higher Score = Worse Symptom Severity.
**Higher Score = Better Function.

ratings of global health change has been reported in other cancer-specific PRO measures (ie, the FACT-G).[23] Despite attenuated responsiveness to retrospectively rated improvements, our study did demonstrate similar positive and negative responsiveness on PROMIS measures across a prospective assessment of change (patient-rated ECOG PS). Given the known methodological concerns using global retrospective change ratings (eg, recall bias, implicit evaluation of changes), our findings suggest prioritizing prospective change assessments in similar validation efforts.

These findings also present evidence that PROMIS measures are responsive to both medical and cancer-specific clinical events. We observed that a hospitalization within this 6-month period, between 6 and 18 months after diagnosis, was linked to clinically meaningful increases

TABLE 4. Longitudinal Change and Effect Sizes by Patient-Reported Outcomes Measurement Information System (PROMIS) Measure

Group Contrast	Score Difference															
	Pain		Fatigue		Anxiety		Depression		Sleep		Physical		Social		Cognitive	
	Mean (ES)	*P*	Mean (ES)	*P*	Mean (ES)	*P*	Mean (ES)	*P*	Mean (ES)	*P*	Mean (ES)	*P*	Mean (ES)	*P*	Mean (ES)	*P*
Full survey																
1-Point ECOG PS improvement (n = 471) vs no change (n = 1868)	−4.5 (0.45)	< .001	−3.2 (0.47)	< .001	−1.8 (0.21)	< .001	−2.2 (0.26)	< .001	−1.2 (0.19)	< .001	3.4 (0.53)	< .001	3.9 (0.49)	< .001	3.0 (0.32)	< .001
1-Point ECOG PS decline (n = 450) vs no change (n = 1868)	3.5 (0.45)	< .001	4.3 (0.63)	< .001	2.8 (0.36)	< .001	3.1 (0.41)	< .001	2.2 (0.37)	< .001	−3.4 (0.62)	< .001	−4.1 (0.57)	< .001	−3.1 (0.4)	< .001
Cancer status self-report: Yes (n = 443) vs remission/no (n = 2008)	2.8 (0.32)	< .001	2.3 (0.31)	< .001	1.9 (0.23)	< .001	1.9 (0.24)	< .001	1.2 (0.20)	.001	−1.9 (−0.30)	< .001	−2.6 (0.32)	< .001	−2.0 (0.23)	< .001
Medical record events																
Hospitalization (n = 38) vs no hospitalization in ≥6 mo (n = 538)	4.7 (0.51)	.002	3.01 (0.35)	.037	2.5 (0.27)	.105	4.79 (0.57)	.001	0.42 (0.06)	.741	−1.78 (−0.28)	.095	−2.35 (0.28)	.093	−0.18 (0.02)	.900
Recurrence/progression (n = 136) vs continued remission (n = 435)	2.34 (0.25)	.015	1.86 (0.22)	.033	0.23 (0.03)	.802	1.05 (0.12)	.232	1.98 (0.27)	.010	−1.21 (0.20)	.059	−1.49 (0.19)	.084	−0.81 (0.09)	.373

Abbreviations: ECOG PS, Eastern Cooperative Oncology Group performance status: ES, effect size.

TABLE 5. Correlations of Selected PROMIS Measures and the FACT-G FWB sub-scale

PROMIS Measure	FACT-G Physical Well-Being Subscale Score		
	Baseline	Follow-Up	Change
Symptoms			
Pain interference	−0.72	−0.71	−0.33
Fatigue	−0.81	−0.82	−0.47
Function			
Physical	0.76	0.75	0.46
Social	0.78	0.78	0.44

Abbreviation: FACT-G, Functional Assessment of Cancer Treatment-General; PROMIS, Patient-Reported Outcomes Measurement Information System.

(range, 3-5 points) in pain, fatigue, and depression. This decline complements recent work examining the responsive of PROMIS measures in surgical recovery after heart transplantation.[7] Small increases in pain, fatigue, and sleep disturbance caused by a documented recurrence/progression of cancer (range, 1.8-2.3 points) provide evidence of responsiveness to cancer-specific clinical events. In contrast, worsening in anxiety and depression was responsive only to patient self-reports of cancer at follow-up and was not based on clinical documentation alone. This difference suggests that anxiety and depression may be more sensitive to patient perception of recurrence or progression rather than clinical identification alone.

Limitations include the long time gap between ratings of change (6 months), which has been identified as creating difficulty in terms of obtaining accurate patient reports.[24] Nevertheless, the association between global ratings of change and actual change scores was reasonably high. MRA comparisons rely on a short period (6 months) to document recurrence/progression and, in some cases (hospitalization), report on very small samples. In addition, no further delineation was made to the 6-month window for evaluating patients after initial treatment, clinical events, or additional cancer-related treatments. Therefore, events could have occurred at any point within this 6-month period, potentially reducing the degree of responsiveness measured by the PROMIS short forms. Further research is necessary, and should focus on tracking the impact of cancer-specific medical events (eg, hospitalizations, adjuvant therapies, recurrence). Finally, it is possible that a scale recalibration-response shift might have occurred over the survey period, resulting in computed change scores that may not have been fully reflective of the true change that has taken place. However, the study was not designed to formally identify or evaluate this occurrence.

These observational data provide a necessary first step in detailing sensitivity to change for 8 PROMIS measures when reported by cancer patients. They lay a ground work for incorporating the PROMIS measures into oncology clinical trials by helping inform sample size needs and power calculations for studies aimed at detecting a specific magnitude of change in 1 of these PRO endpoints. In addition, the data can help interpret the magnitude of change or differences in results observed with 1 or more PROMIS measures. These data also are timely, given the increased interest by the US Food and Drug Administration (FDA) in patient-focused drug development and the release of the FDA *Clinical Outcomes Assessment Compendium* to help guide applications to their Clinical Outcomes Assessment Qualification Program, as industry adoption of PROs increases.[25]

Finally, although we have many cancer-specific PRO measures to choose from for use in cancer trials, few generic PRO measures have been validated with an appropriate degree of responsiveness in samples of patients who have cancer. These findings are among the first to indicate that PROMIS measures may be able to compare the magnitude of benefit or harm from treatments across both cancer and noncancer clinical trials (eg, for pain or symptom management studies or for measures of clinical benefit). This provides an opportunity to make meaningful comparisons across heath conditions and a broader context for the interpretation of PRO endpoints.

In conclusion, the current study presents strong evidence across multiple evaluation methods that PROMIS measures are responsive to both improvements and declines in symptoms and function experienced by patients with cancer. It extends past work, presenting further evidence that clinically meaningful changes across PROMIS measures range from 3 to 5 points. These results also highlight the utility using PROMIS measures in research (clinical trials, observational cohorts, and comparative effectiveness evaluations).

FUNDING SUPPORT

This work was supported by grants from the National Institutes of Health (NIH) (U01 AR057971 [to R.E.J., T.L., A.L.P, and C.M.M.] and P30 CA051008 [to R.E.J., T.L., and A.L.P.]); the National Center for Research Resources, the National Center for Advancing Translational Sciences, and the NIH through the Clinical and Translational Science Awards Program (KL2 TR000102 [to R.E.J.])

CONFLICT OF INTEREST DISCLOSURES

David Cella serves on the PROMIS Health Organization Board of Directors.

AUTHOR CONTRIBUTIONS

Roxanne E. Jensen: Conception and design, acquisition and assembly of data, and data analysis and interpretation. **Carol M. Moinpour**: Conception and design, financial support, acquisition and assembly of data, and data analysis and interpretation. **Arnold L. Potosky**: Conception and design, financial support, acquisition and assembly of data, and data analysis and interpretation. **Tania Lobo**: Acquisition and assembly of data and data analysis and interpretation. **Elizabeth A. Hahn**: Data analysis and interpretation **Ron D. Hays**: Conception and design. **David Cella**: Conception and design and data analysis and interpretation. **Ashley Wilder Smith**: Conception and design, financial support, data analysis and interpretation. **Xiao-Cheng Wu**: Acquisition and assembly of data. **Theresa H. M. Keegan**: Acquisition and assembly of data. **Lisa E. Paddock**: Acquisition and assembly of data. **Antoinette M. Stroup**: Acquisition and assembly of data. **David T. Eton**: Data analysis and interpretation. All authors wrote and approved the final article and were accountable for all aspects of the work.

REFERENCES

1. US Food and Drug Administration (FDA), Department of Health and Human Services. Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Silver Spring, MD: US Department of Health and Human Services; 2009.
2. National Institute of Environmental Health Sciences. NIH Roadmap and Roadmap-affiliated Initiatives. Available at: https://commonfund.nih.gov/grants/fundedresearch. Accessed September 9, 2016.
3. The Patient-Reported Outcomes Measurement Information System (PROMIS). PROMIS Overview. Available at: http://www.nihpromis.org/about/overview. Accessed September 9, 2016.
4. Hays RD, Hadorn D. Responsiveness to change: an aspect of validity, not a separate dimension. *Qual Life Res*. 1992;1:73-75.
5. Hays RD, Spritzer KL, Fries JF, Krishnan E. Responsiveness and minimally important difference for the Patient-Reported Outcomes Measurement Information System (PROMIS) 20-item physical functioning short form in a prospective observational study of rheumatoid arthritis. *Ann Rheum Dis*. 2015;74:104-107.
6. Pilkonis PA, Yu L, Dodds NE, Johnston KL, Maihoefer CC, Lawrence SM. Validation of the depression item bank from the Patient-Reported Outcomes Measurement Information System (PROMIS) in a 3-month observational study. *J Psychiatr Res*. 2014; 56:112-119.
7. Flynn KE, Dew MA, Lin L, et al. Reliability and construct validity of PROMIS measures for patients with heart failure who undergo heart transplant. *Qual Life Res*. 2015;24:2591-2599.
8. Yost KJ, Eton DT, Garcia SF, Cella D. Minimally important differences were estimated for 6 Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. *J Clin Epidemiol*. 2011;64:507-516.
9. Jensen RE, Moinpour CM, Keegan THM, et al. The Measuring Your Health Study: leveraging community-based cancer registry recruitment to establish a large, diverse cohort of cancer survivors for analyses of measurement equivalence and validity of the Patient Reported Outcomes Measurement Information System (PROMIS) short form items. *Psychol Test Assess Model*. 2016;58:99-117.
10. Lohr KN. Health Outcomes Methodology Symposium: summary and recommendations. *Med Care*. 2000;38:II-1–II-210.
11. Institute of Medicine, Smedley B, Stith A, Nelson A, eds. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Washington, DC: The National Academies Press; 2002.
12. Jensen RE, Potosky AL, Reeve BB, et al. Validation of the PROMIS physical function measures in a diverse US population-based cohort of cancer patients. *Qual Life Res*. 2015;24:333-2344.
13. US Census Bureau, US Department of Commerce, Economics and Statistics Administration. Overview of Race and Hispanic Origin:

2010. 2010 Census Briefs. Issued March 2011. Available at: http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf.
14. Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *J Clin Oncol*. 1993;11:570-579.
15. Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol*. 1982;5:649-655.
16. 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;63:1179-1194.
17. Lai JS, Wagner LI, Jacobsen PB, Cella D. Self-reported cognitive concerns and abilities: 2 sides of 1 coin? *Psychooncology*. 2014;23:1133-1141.
18. Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. *Eval Health Prof*. 2005;28:172-191.
19. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
20. Cella D, Eton DT, Fairclough DL, et al. What is a clinically meaningful change on the Functional Assessment of Cancer Therapy-Lung (FACT-L) questionnaire? Results from Eastern Cooperative Oncology Group (ECOG) Study 5592. *J Clin Epidemiol*. 2002;55:285-295.
21. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61:102-109.
22. Cook KF, Jensen SE, Schalet BD, et al. PROMIS measures of pain, fatigue, negative affect, physical function, and social function demonstrated clinical validity across a range of chronic conditions. *J Clin Epidemiol*. 2016;73:89-102.
23. Cella D, Hahn EA, Dineen K. Meaningful change in cancer-specific quality of life scores: differences between improvement and worsening. *Qual Life Res*. 2002;11:207-221.
24. Stull DE, Leidy NK, Parasuraman B, Chassany O. Optimal recall periods for patient-reported outcomes: challenges and potential solutions. *Curr Med Res Opin*. 2009;25:929-942.
25. US Food and Drug Administration (FDA). Clinical Outcome Assessment Compendium, 2015. Silver Spring, MD: US Department of Health and Human Services; 2015.