Published in final edited form as:

Qual Life Res. 2014 September ; 23(7): 2031–2035. doi:10.1007/s11136-014-0647-2.

Comparing PROMIS computer adaptive tests to the Brief Symptom Inventory in patients with prostate cancer

George Baum,

Department of Behavioral Science, The University of Texas, MD Anderson Cancer Center P.O. Box 301439, Unit 1330 Houston, TX 77230-1439

Karen Basen-Engquist,

Department of Behavioral Science, The University of Texas, MD Anderson Cancer Center P.O. Box 301439, Unit 1330 Houston, TX 77230-1439

Maria C. Swartz,

Department of Behavioral Science, The University of Texas, MD Anderson Cancer Center P.O. Box 301439, Unit 1330 Houston, TX 77230-1439

The University of Texas School of Public Health at Houston P.O. Box 20186 Houston, TX 77025

Patricia A. Parker, and

Department of Behavioral Science, The University of Texas, MD Anderson Cancer Center P.O. Box 301439, Unit 1330 Houston, TX 77230-1439

Cindy L. Carmack

Department of Behavioral Science, The University of Texas, MD Anderson Cancer Center P.O. Box 301439, Unit 1330 Houston, TX 77230-1439

Abstract

Purpose—This study assessed whether the Patient-Reported Outcomes Measurement Information System (PROMIS) computer adaptive tests (CATs) provided results similar to those of the Brief Symptom Inventory (BSI) with a low patient burden.

Methods—Secondary data analysis of 136 prostate cancer patients who completed the 53-item BSI and the PROMIS CATs assessing depression, anxiety, and hostility.

Results—The PROMIS CATs and BSI correlated significantly in measures of depression (.85), anxiety (.76), and anger/hostility (.66; $p < .001$ for all). Using our BSI cut-off points for depression, anxiety, and anger/hostility, ROC analysis yielded areas under the curve of .966 (standard error [SE] = .014, $p < .001$), .975 (SE = .012, $p < .001$), and .952 (SE = .027, $p < .001$), respectively.

Conclusions—PROMIS CAT tests were highly correlated with the BSI subscales, indicating that the CATs performed well compared with the BSI, a widely used psychosocial measure.

Keywords

Computer adaptive tests; BSI; Brief Symptom Inventory; Prostate cancer; Quality of life; PROMIS; Health surveys

Introduction

In the United States, approximately 241,740 men were diagnosed with prostate cancer in 2012 and 28,170 died from the disease [1]. Prostate cancer patients often experience declining physical functioning and other problems that affect quality of life (QOL) and increase psychological distress [2]. Prostate cancer survivors' psychological functioning is markedly understudied [3], available studies, however, indicate that for some, post cancer treatment psychological adjustment declines over time [4].

In 2004 the National Institutes of Health (NIH) began the "Roadmap Initiative" for medical research and part of this program was the Patient-Reported Outcomes (PROs) Measurement Information System (PROMIS) [5]. The purpose of PROMIS was to improve the reliability and validity of PROs and to develop new instruments that could improve outcome measures for various chronic diseases [6]. PROMIS identified important PRO domains and developed item banks for the following six: physical functioning, fatigue, pain, emotional distress, social role participation, sleep disturbance, and wake disturbance [5, 7].

There are many existing standardized questionnaires to assess these PRO domains, but many are lengthy, and their results are often missing data because patients may not want to complete or feel up to completing all of the questionnaire items. Computer adaptive tests (CATs) adapt to the respondent by sequentially choosing only the most informative items for an individual depending on his or her earlier responses. With CATs, scores can be obtained by administering fewer items [8], and testing can be tailored to each patient. Furthermore, lengthy standardized questionnaires often include items that cover the entire continuum of the PRO domain in question, whereas CATs use an algorithm to reduce the number of items necessary to determine the subject's level on the PRO continuum [8, 9]. Traditionally, CATs have not been widely used in medical research, but with the advent of the PROMIS initiative, CAT usage started becoming more prevalent [10-12].

The CATs tools need to be validated across a range of cancer populations to determine whether they will be useful in clinical research and whether they reduce respondent burden. The purpose of this study was to compare the PROMIS CATs with the validated, 53-item Brief Symptom Inventory (BSI) [13, 14] in a sample of non-metastatic prostate cancer patients. Another purpose of this study was to assess the PROMIS CATs' performance when using cut-off points for psychological distress.

Methods

We conducted a secondary analysis using data that were collected during the screening of prostate cancer patients recruited to a psychosocial intervention study. Briefly, that study was a two-condition randomized pilot test to determine whether distressed patients benefited

more from a psychosocial group intervention if it was heterogeneous (i.e., composed of distressed and non-distressed members) or if it was homogeneous (i.e., distressed members only) [15]. Eligible participants were given a screening assessment that included the BSI and the PROMIS CATs that assessed depression, anxiety, and anger.

The BSI yields both continuous raw score values and adjusted area T scores with a mean of 50 and a SD of 10 for each of nine symptom subscales with higher scores indicating higher distress. We used the adult nonpatient norms for males and females for these analyses.

“Caseness” on the BSI is defined by a T score of ≥ 63 on two of its nine subscales or on its Global Severity Index [16]. Since there are no established cut-off points indicating caseness for the individual subscales, we used a T score of ≥ 63 as our cut-off point for each subscale. Using a reference population of the U.S. general population, PROMIS CAT scales produce continuous T scores with a mean of 50 and a SD of 10, higher scores indicating higher distress [17]. The BSI has 6 items in the anxiety subscale, 5 items in the hostility subscale and 6 items in the depression subscale giving a total of 17 items for those three domains. The PROMIS CAT tests varied in the number of items based on the answers the respondent gave.

To compare the PROMIS CAT depression, anxiety, and anger scales with the BSI depression, anxiety, and hostility subscales, the following analyses were conducted. Bivariate correlations were performed between each BSI subscale and the respective PROMIS CAT scale. Receiver operating characteristic curve (ROC) analysis was performed on the PROMIS CAT depression, anxiety, and anger scales by using the BSI cut-off points as the binary classifier.

Results

Our analyses included 136 prostate cancer patients who were 65% white and were a mean age of 64.5 years (Table 1) all had non-metastatic cancer. The mean T scores standard deviation (SD) for the BSI scales were 52.3 ± 10.6 for depression, 50.7 ± 11.2 for anxiety, and 50.2 ± 10.1 for hostility, indicating that this population had levels of depression, anxiety and hostility that were in line with the referent population from the BSI. The mean scores for the PROMIS CAT measures were 45.9 ± 8.9 for depression, 48.6 ± 8.2 for anxiety, and 47.2 ± 8.2 for anger, indicating that these subjects had slightly lower levels of these three domains than the general population on which the T scores were based. The average numbers of items administered by the PROMIS CATs were 9.5 ± 6.5 items for depression, 7.4 ± 4.3 items for anxiety, and 7.1 ± 2.3 items for anger.

Correlations between the PROMIS CAT and the BSI depression, anxiety, and anger/hostility scales were .85 ($p < .001$), .76 ($p < .001$) and .66 ($p < .001$) respectively (Table 2).

ROC analysis

Using our BSI cut-off point for depression, the ROC analysis of the PROMIS depression scale yielded an area under the curve of .966 (standard error [SE] = .014, $p < .001$) (Table 2). A cut-off point of 53.1 on the PROMIS CAT depression scale yielded a specificity of .95 and a sensitivity of .81 (Table 3). Using our BSI cut-off point for anxiety, the area under the

curve for the PROMIS CAT anxiety scale was .975 (SE = .012, $p < .001$) (Table 2) with a specificity of .96 and sensitivity of .84 at a cut-off point of 56.1 on the PROMIS CAT anxiety scale (Table 3). When the PROMIS CAT anger scale was given a cut-off of 54.5 (Table 3), the area under the curve was .952 (SE = .027, $p < .001$) with our BSI hostility cut-off point (Table 2 and Fig. 1). Table 3 shows a few other potential cut-off points for the PROMIS scales and their corresponding sensitivity and specificity values based on the results from the ROC analyses.

Discussion

The correlations between the PROMIS CAT and BSI depression, anxiety, and anger/hostility scales were high, and the ROC results demonstrated a large area under the curve using our BSI cut-off points. This finding is consistent with earlier research that found high correlations between PROMIS tools and other legacy instruments when comparing similar domains [17, 18].

One of the hypothesized benefits of CATs is that they reduce respondent burden [6]. In our study, the participants answered 24 items on average to complete the PROMIS CATs for depression, anxiety, and anger combined. The BSI is a 53-item instrument, with 17 total items for depression, anxiety, and hostility. Hence, the PROMIS CAT scales resulted in a slightly higher respondent burden than the corresponding BSI subscales did. However, another potential benefit of CAT tests is the possibility of increased precision. CATs select questions for the purpose of maximizing precision and theoretically have higher precision than static fixed instruments [19]. The PROMIS CAT tests are derived from large item banks and therefore may be more precise than the 5 or 6 items used for the different dimensions on the BSI.

This study adds to the previous research by comparing the PROMIS CAT tools to validated PRO instruments. Other studies have used PROMIS CATs in cancer populations [20], and PROMIS CATs have been successfully validated in other disease-specific populations [21-23]. Like these studies, our study demonstrated a strong correlation between the PROMIS CATs and validated legacy instruments, strengthening the argument that PROMIS CATs are a viable alternative to traditional PRO measures.

Cut-off scores could be important for clinical and research purposes; thus, we explored potential cut-off points for depression, anxiety, and anger by using the PROMIS CAT scales. Using a BSI cut-off point of a T score of 63, the PROMIS CAT cut off points scores of 53.1, 56.1, and 54.5 for the depression, anxiety, and anger scales, respectively, yielded high specificity and sensitivity. However, the PROMIS scales were not developed as a diagnostic tool to determine caseness in a clinical setting, so at this point these results are purely exploratory.

Our findings need to be interpreted with caution. The sample was relatively small and included only patients with prostate cancer. In addition, these cancer patients were, on average, older than the general population. Furthermore, the BSI cut-off points used in this study may not be the best “gold standard” to test the function of the PROMIS CATs,

especially when comparing the PROMIS CAT anger scale with the BSI hostility scale, because these two domains may not map to each other as well as the PROMIS and BSI depression scales or the PROMIS and BSI anxiety scales, as indicated by their correlation.

Finally, an issue that researchers have with using CATs is the fact that participants are given different items to arrive at a final score; also, the same participant could answer different items over time in a longitudinal study. Hence, future research should also focus on comparing PROMIS CATs with other validated tools over time and in trial settings. CATs are derived using Item Response Theory (IRT) and one of the advantages of this approach is greater precision with fewer items administered [19]. So, despite the weaknesses we have pointed out, the potential for greater precision with lower response burden makes it important to test the PROMIS CATs in various populations. CATs are widely used in many fields and have the potential to benefit medical researchers.

Acknowledgements

This study is supported in part by the Patient-Reported Outcomes, Survey, and Population Research (PROSPR) Shared Resource; the Cancer Center Support Grant CA16672, given to The University of Texas MD Anderson Cancer Center by the National Institutes of Health; and 5R21CA126854-02, given to Cindy L. Carmack by the National Cancer Institute.

Abbreviations

PRO	Patient-reported outcome
CAT	Computer adaptive test
QOL	Quality of life
NIH	National Institutes of Health
PROMIS	Patient Reported Outcomes Measurement Information System
BSI	Brief Symptom Inventory
ROC	Receiver operating characteristic curve

References

1. American Cancer Society. Cancer Facts & Figures. American Cancer Society; Atlanta, GA: 2012. p. 2012
2. da Silva FC, et al. Quality of life in patients with prostatic cancer. A feasibility study. The Members of Quality of Life Committee of the EORTC Genitourinary Group. Cancer. 1993; 71(3 Suppl): 1138–42. [PubMed: 8428336]
3. Kornblith A. Psychosocial adaptation of cancer survivors. Psycho-oncology. 1998; 1:223–254.
4. Eton DT, Lepore SJ. Prostate cancer and health-related quality of life: a review of the literature. Psycho-oncology. 2002; 11(4):307–326. [PubMed: 12203744]
5. National Institutes of Health. [2014 January 4th] PROMIS: Patient-Reported Outcomes Measurement Information System.. Current Programs. 2013. Available from: <https://commonfund.nih.gov/promis/overview>
6. Fries JF, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. Clin Exp Rheumatol. 2005; 23(5 Suppl 39):S53–7. [PubMed: 16273785]

7. Reeve BB, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care*. 2007; 45(5 Suppl 1):S22–31. [PubMed: 17443115]
8. Fliege H, et al. Development of a Computer-adaptive Test for Depression (D-CAT). *Quality of Life Research*. 2005; 14(10):2277–2291. [PubMed: 16328907]
9. Cheville AL, et al. Performance of an Item Response Theory-Based Computer Adaptive Test in Identifying Functional Decline. *Archives of Physical Medicine and Rehabilitation*. 2012; 93(7): 1153–1160. [PubMed: 22749314]
10. Carle AC, et al. Advancing PROMIS's methodology: results of the Third Patient-Reported Outcomes Measurement Information System (PROMIS®) Psychometric Summit. *Expert Review of Pharmacoeconomics & Outcomes Research*. 2011; 11(6):677–684. [PubMed: 22098283]
11. Gershon RC, et al. The use of PROMIS and assessment center to deliver patient-reported outcome measures in clinical research. *Journal of applied measurement*. 2010; 11(3):304. [PubMed: 20847477]
12. Rothrock NE, 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). *Journal of Clinical Epidemiology*. 2010; 63(11):1195–1204. [PubMed: 20688471]
13. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. *Psychol Med*. 1983; 13(3):595–605. [PubMed: 6622612]
14. Herr HW, Kornblith AB, Ofman U. A comparison of the quality of life of patients with metastatic prostate cancer who received or did not receive hormonal therapy. *Cancer*. 1993; 71(3 Suppl): 1143–50. [PubMed: 8428337]
15. Carmack Taylor CL, et al. A social comparison theory analysis of group composition and efficacy of cancer support group programs. *Social Science & Medicine*. 2007; 65(2):262–273. [PubMed: 17448580]
16. Derogatis, LR. BSI, Brief Symptom Inventory: administration, scoring & procedures manual. National Computer Systems; 1993.
17. Pilkonis PA, et al. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS(R)): depression, anxiety, and anger. *Assessment*. 2011; 18(3):263–83. [PubMed: 21697139]
18. Cella D, 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):1179–94. [PubMed: 20685078]
19. Weiss DJ. Adaptive testing by computer. *Journal of consulting and clinical psychology*. 1985; 53(6):774. [PubMed: 3841355]
20. Yost KJ, et al. Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. *Journal of Clinical Epidemiology*. 2011; 64(5):507–516. [PubMed: 21447427]
21. Khanna D, et al. Patient-Reported Outcomes Measurement Information System (PROMIS(R)) -- The future of measuring patient reported outcomes in rheumatology. *Arthritis Care Res (Hoboken)*. 2011; 63(S11):S486–S490.
22. Khanna D, et al. Feasibility and construct validity of PROMIS and “legacy” instruments in an academic scleroderma clinic. *Value Health*. 2012; 15(1):128–34. [PubMed: 22264980]
23. Spiegel BM, et al. Understanding gastrointestinal distress: a framework for clinical practice. *Am J Gastroenterol*. 2011; 106(3):380–5. [PubMed: 21378758]

a PROMIS CAT depression curve **b** PROMIS CAT anxiety curve

c PROMIS CAT anger curve

Fig. 1.
ROC curves for PROMIS CAT scales for depression, anxiety and anger using corresponding
BSI cut-off points

Table 1Demographic characteristics of the sample ($n = 136$)

Characteristic	
Race, n (%)	
White	89 (65%)
African American	31 (23%)
Hispanic	11 (8%)
Asian	4 (3%)
Other	1 (1%)
Age, mean (SD)	64.5 (7.8) years

Table 2

Correlations between BSI and PROMIS CAT scales and ROC analyses of PROMIS CAT scales using BSI cut-offs

Domain	Correlation (<i>p</i> value)	ROC area under curve (SE, <i>p</i> value)
Depression	.85 (< .001)	.966 (.014, < .001)
Anxiety	.76 (< .001)	.975 (.012, < .001)
Anger/Hostility	.66 (< .001)	.952 (.027, < .001)

Table 3

PROMIS CAT sensitivity and specificity at various cut-off points

PROMIS CAT cut-off scores	Sensitivity	Specificity	SD from mean
Depression scale			
49.5	.96	.86	.40
53.1	.81	.95	.81
57.8	.50	.99	1.34
Anxiety scale			
54.7	.95	.93	.74
56.1	.84	.96	.91
61.5	.47	.99	1.57
Hostility scale			
51.5	.993	.821	.52
54.5	.800	.949	.89
58.1	.667	.983	1.33