



Published in final edited form as:

*JAMA Oncol.* 2015 November ; 1(8): 1051–1059. doi:10.1001/jamaoncol.2015.2639.

## Validity and Reliability of the U.S. National Cancer Institute's Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE)

**Amylou C. Dueck, PhD,**

Mayo Clinic, Scottsdale, AZ, USA

**Tito R. Mendoza, PhD,**

University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

**Sandra A. Mitchell, PhD, CRNP, AOCN,**

National Cancer Institute, Rockville, MD, USA

**Bryce B. Reeve, PhD,**

University of North Carolina, Chapel Hill, NC, USA

**Kathleen M. Castro, RN, MS, AOCN,**

National Cancer Institute, Rockville, MD, USA

**Lauren J. Rogak, MA,**

Memorial Sloan Kettering Cancer Center, New York, NY, USA

**Thomas M. Atkinson, PhD,**

Memorial Sloan Kettering Cancer Center, New York, NY, USA

**Antonia V. Bennett, PhD,**

University of North Carolina, Chapel Hill, NC, USA

**Andrea M. Denicoff, MS, RN, ANP,**

National Cancer Institute, Rockville, MD, USA

---

Corresponding author: Amylou C. Dueck, Mayo Clinic, 13400 E. Shea Blvd, Scottsdale, AZ 85259, (480) 301-6159, dueck@mayo.edu.

\*Additional National Cancer Institute PRO-CTCAE Study Group members are: Narre Heon, Mary Shaw, Sean Ryan, Liora P. Stark, Laura Sit, Amy Thomassie, Wendy Pettus, Lucy Gansauer, Barbara Marconi, Angela Steele-Tilton, Donna Malveaux, Jennifer Wind  
Narre Heon, BA (Memorial Sloan Kettering Cancer Center, New York, NY, USA) heonn@mskcc.org  
Mary Shaw, BA (Memorial Sloan Kettering Cancer Center, New York, NY, USA) ShawM@mskcc.org  
Sean Ryan, MS (Memorial Sloan Kettering Cancer Center, New York, NY, USA) ryans@mskcc.org  
Liora P. Stark, MPH (Memorial Sloan Kettering Cancer Center, New York, NY, USA) Liorastark@gmail.com  
Laura Sit, MS (Memorial Sloan Kettering Cancer Center, New York, NY, USA) laura.sit@gmail.com  
Amy Thomassie RN, BSN, CCRC (The Cancer Program of Our Lady of the Lake and Mary Bird Perkins, Baton Rouge, LA, USA) Amy.Thomassie@pbrc.edu  
Wendy Pettus (Gibbs Cancer Center, Spartanburg, SC, USA) WPettus@srhs.com  
Lucy Gansauer, RN, MSN, OCN (Gibbs Cancer Center, Spartanburg, SC, USA) lgansauer@gibbscc.org  
Barbara Marconi, RN (Helen F. Graham Cancer Center & Research Institute, Christiana Care Health System, Newark, DE, USA) BMarconi@Christianacare.org  
Angela Steele-Tilton, MSN (Helen F. Graham Cancer Center & Research Institute, Christiana Care Health System, Newark, DE, USA) Angela.b.steele-tilton@wilmu.edu  
Donna Malveaux (University of Texas M. D. Anderson Cancer Center, Houston, TX, USA) dmalveau@mdanderson.org  
Jennifer Wind, MA (Dana-Farber Cancer Institute, Boston, MA, USA) Jennifer\_Wind@DFCI.HARVARD.EDU

**Trial Registration:** ClinicalTrials.gov NCT02158637; <http://www.clinicaltrials.gov/ct2/show/NCT02158637>

**Ann M. O'Mara, PhD, RN, FAAN,**  
National Cancer Institute, Rockville, MD, USA

**Yuelin Li, PhD,**  
Memorial Sloan Kettering Cancer Center, New York, NY, USA

**Steven B. Clauser, PhD, MPA,**  
National Cancer Institute, Rockville, MD, USA

**Donna M. Bryant, MSN, ANP-BC, OCN, CCRC,**  
The Cancer Program of Our Lady of the Lake and Mary Bird Perkins, Baton Rouge, LA, USA

**James D. Bearden, MD, FACP,**  
Gibbs Cancer Center, Spartanburg, SC, USA

**Theresa A. Gillis, MD,**  
Helen F. Graham Cancer Center & Research Institute, Christiana Care Health System, Newark, DE, USA

**Jay K. Harness, MD,**  
The Center for Cancer Prevention and Treatment, St. Joseph Hospital of Orange, Orange, CA, USA

**Robert D. Siegel, MD, FACP,**  
Hartford Hospital-Helen and Harry Gray Cancer Center, Hartford, CT, USA

**Diane B. Paul [Patient Advocate and Cancer Survivor],**  
Brooklyn, NY, USA

**Charles S. Cleeland, PhD,**  
University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

**Deborah Schrag, MD, MPH,**  
Dana-Farber Cancer Institute, Boston, MA, USA

**Jeff A. Sloan, PhD,**  
Mayo Clinic, Rochester, MN, USA

**Amy P. Abernethy, MD, PhD,**  
Duke University Medical Center, Durham, NC, USA

**Deborah W. Bruner, RN, PhD, FAAN,**  
Emory University School of Nursing, Atlanta, GA, USA

**Lori M. Minasian, MD, FACP,**  
National Cancer Institute, Rockville, MD, USA

**Ethan Basch, MD, MSc, and**  
University of North Carolina, Chapel Hill, NC, USA

**on behalf of the National Cancer Institute PRO-CTCAE Study Group\***

Tito R. Mendoza: tmendoza@mdanderson.org; Sandra A. Mitchell: mitchlls@mail.nih.gov; Bryce B. Reeve: bbreeve@email.unc.edu; Kathleen M. Castro: kathleen.castro@nih.gov; Lauren J. Rogak: rogakl@mskcc.org; Thomas M. Atkinson: atkinsot@mskcc.org; Antonia V. Bennett: avbenn@unc.edu; Andrea M. Denicoff: denicofa@mail.nih.gov; Ann M. O'Mara: omaraa@mail.nih.gov; Yuelin Li: LiY12@mskcc.org; Steven B. Clauser: sclouser@pcori.org; Donna M. Bryant: dbryantnp@cox.net; James D. Bearden: jbearden@srhs.com; Theresa A. Gillis: TGillis@christianacare.org; Jay K.

Harness: jkharness@gmail.com; Robert D. Siegel: ROBERT\_SIEGEL@bshsi.org; Diane B. Paul: funnylady93@nyc.rr.com; Charles S. Cleeland: ccleeland@mdanderson.org; Deborah Schrag: deb\_schrag@dfci.harvard.edu; Jeff A. Sloan: jsloan@mayo.edu; Amy P. Abernethy: amy.abernethy@duke.edu; Deborah W. Bruner: deborah.w.bruner@emory.edu; Lori M. Minasian: minasilo@mail.nih.gov; Ethan Basch: ebasch@med.unc.edu

## Abstract

**Importance**—Symptomatic adverse events (AEs) in cancer trials are currently reported by clinicians using the National Cancer Institute's (NCI) Common Terminology Criteria for Adverse Events (CTCAE). To integrate the patient perspective, the NCI developed a patient-reported outcomes version of the CTCAE (PRO-CTCAE) to capture symptomatic AEs directly from patients.

**Objective**—To assess the construct validity, test-retest reliability, and responsiveness of PRO-CTCAE items.

**Design**—Participants completed PRO-CTCAE items on tablet computers in clinic waiting rooms at two visits 1-6 weeks apart. A subset completed PRO-CTCAE items during an additional visit one business day after the first visit.

**Setting**—Nine U.S. cancer centers and community oncology practices.

**Participants**—975 adult cancer patients undergoing outpatient chemotherapy and/or radiation enrolled between January 2011 and February 2012. Eligibility required participants to read English and be without clinically significant cognitive impairment.

**Main Outcome(s) and Measure(s)**—Primary comparators were clinician-reported Eastern Cooperative Oncology Group Performance Status (ECOG PS) and the European Organisation for Research and Treatment of Cancer Core Quality of Life Questionnaire (QLQ-C30).

**Results**—940/975 (96%) and 852/940 (91%) participants completed PRO-CTCAE items at each visit. 938/940 (99.8%) participants (53% female, median age 59, 32% high school education or less, 17% ECOG PS 2-4) reported having at least one symptom. All PRO-CTCAE items had at least one correlation in the expected direction with a QLQ-C30 scale (111/124  $P < .05$ ). Stronger correlations were seen between PRO-CTCAE items and conceptually-related QLQ-C30 domains. Scores for 94/124 PRO-CTCAE items were higher in the ECOG PS 2-4 versus 0-1 group (58/124  $P < .05$ ). Overall, 119/124 items met at least one construct validity criterion. Test-retest reliability was acceptable for 36/49 pre-specified items (median intra-class correlation coefficient .76; range .53-.96). Correlations between PRO-CTCAE item changes and corresponding QLQ-C30 scale changes reached statistical significance for 27 pre-specified items (median  $r = .43$ , range .10-.56; all  $P < .006$ ).

**Conclusions and Relevance**—Evidence demonstrates favorable validity, reliability, and responsiveness of PRO-CTCAE in a large, heterogeneous U.S. sample of patients undergoing cancer treatment. Studies evaluating other measurement properties of PRO-CTCAE are underway to inform further development of PRO-CTCAE and its inclusion in cancer trials.

## Introduction

In cancer clinical trials, adverse events (AEs) are collected and reported using the U.S. National Cancer Institute's (NCI) Common Terminology Criteria for Adverse Events

(CTCAE).<sup>1</sup> The CTCAE is a library of items representing 790 discrete AEs, each graded using an ordinal severity scale.<sup>2</sup> Approximately 10% of AEs in the CTCAE are symptoms (e.g., nausea, sensory neuropathy), which in trials have historically been reported by clinical investigators.<sup>3</sup> However, there is empiric evidence that collection of this information directly from patients improves the precision and reliability of symptomatic AE detection in trials,<sup>4-9</sup> and is feasible.<sup>10,11</sup> Moreover, there is substantial evidence that clinical investigators may miss up to half of patients' symptomatic AEs.<sup>5,6,12,13</sup>

To improve precision and patient-centeredness in the capture of symptomatic AEs, the NCI developed a library of patient-reported outcome (PRO) items to supplement the CTCAE, called the PRO-CTCAE,<sup>14</sup> as has been previously described.<sup>15</sup> Of the 790 AEs in the CTCAE, 78 were identified as amenable to patient self-report. For each of these AEs, PRO items were created reflecting the attributes of frequency, severity, interference with usual or daily activities, amount, or presence/absence. One to three attributes were selected for any given AE depending on the content of the CTCAE criteria for that AE and the nature of that particular AE. In total, 124 individual items represent the 78 symptomatic AEs currently in the PRO-CTCAE item library.

The generic structure for PRO-CTCAE items and response options are shown in Table 1. Each item includes a plain language term for the AE, the attribute of interest, and the standard recall period of “the past 7 days”. Cognitive interviews previously determined a high level of patient understanding and meaningfulness of the items.<sup>16</sup> Software was developed for administering PRO-CTCAE items to patients either via web or an automated telephone interactive voice response (IVR) interface, and was refined through usability testing.<sup>15,17</sup>

For any new measurement tool in clinical research (e.g., biomarkers, imaging, diagnostic test), it is essential to establish that the new instrument accurately and reliably captures the underlying phenomenon it is intended to measure. To accomplish this for the PRO-CTCAE, this study was designed to evaluate the measurement properties of the 124 items in the PRO-CTCAE item library including validity (degree to which an instrument accurately measures the underlying phenomenon), reliability (ability of an instrument to produce similar scores on repeated measurements under similar conditions), and responsiveness (capacity of an instrument to show a change when there has been a change in the underlying phenomenon). These properties were examined individually for each item since PRO-CTCAE items are individually reported in trials and not aggregated into a single score. Inclusion of patients with diversity with respect to cancer type, treatment modality, and sociodemographic characteristics was considered essential given the intended use of PRO-CTCAE across varying research contexts. To simultaneously evaluate the measurement properties of 124 items within a single study required us to employ a varied set of comparators or “anchors”, and warranted a larger and more diverse sample of respondents and settings than is typically employed in most validation studies of fixed length PRO measures.

## Methods

### Patients

Adult patients initiating or undergoing outpatient chemotherapy, radiation, or both at one of nine U.S. cancer centers or community oncology practices were approached in clinical waiting areas and invited to participate in this study. Participating sites with number of patients enrolled included Dana-Farber Cancer Institute, Boston, MA (N=40); Hartford Hospital-Helen and Harry Gray Cancer Center, Hartford, CT (N=104); Helen F. Graham Cancer Center & Research Institute at Christiana Care Health System, Newark, DE (N=105); Mayo Clinic, Rochester, MN (N=9); Memorial Sloan Kettering Cancer Center, New York, NY (N=280); Our Lady of the Lake and Mary Bird Perkins Cancer Center, Baton Rouge, LA (N=133); Gibbs Cancer Center, Spartanburg, SC (N=113); St. Joseph Hospital of Orange, Orange, CA (N=104); and University of Texas M. D. Anderson Cancer Center, Houston, TX (N=52).

Eligibility criteria required that all participants be able to read and comprehend English, be without clinically significant cognitive impairment based on site investigator judgment, have a cancer diagnosis, and be actively undergoing cancer treatment or be initiating treatment in the next 7 days. Patients with any cancer type were eligible, but an accrual strategy was used to enrich for specific cancer types in order to facilitate planned comparisons between groups based on cancer type in the validity analysis, including breast; aerodigestive tract (head/neck and esophageal cancer); genitourinary (prostate and bladder); lung; colorectal; and lymphoma/myeloma. An enrichment strategy was also employed to ensure that a minimum of 15% of participants had impaired performance status (PS) defined as Eastern Cooperative Oncology Group (ECOG) PS  $\geq 2$ .

Study sites were selected to encompass geographic, racial/ethnic, economic, and educational diversity reflective of the U.S. population with the understanding that the requirement to be English speaking would limit the enrollment of Hispanic patients (a separate study evaluating the Spanish language version of the PRO-CTCAE has been conducted<sup>18</sup>). Race/ethnicity was self-reported by patients.

Institutional review board approval was obtained at all sites and at the NCI, and all patients completed written informed consent. The trial was registered on ClinicalTrials.gov (NCT02158637). Each participant received a \$20 gift card or parking voucher.

### Questionnaire

The previously developed PRO-CTCAE item library consists of 78 symptomatic AEs represented by 124 distinct items.<sup>14,15</sup> To limit burden, a maximum of 58 symptomatic AEs (82 items) was presented to each participant. Seven electronic surveys targeted towards different cancer types (eTable 1) were created in the central PRO-CTCAE web survey administration platform. As part of the registration process, the site coordinator selected a single survey based on the patient's diagnosis, and that survey was then automatically scheduled for completion at each visit. All surveys included a set of 20 "core" symptomatic AEs<sup>15</sup>, predetermined based on high prevalence across cancer types in prior NCI-sponsored clinical trials.<sup>19</sup> Remaining symptomatic AEs were classified a priori as likely to be

prevalent or non-prevalent in specific cancer types based on expert consultation, patient representative input, and literature review. These items were included on surveys for selected cancer types to facilitate planned comparisons between groups based on cancer type. When 80% of accrual was reached, to increase sample size for the 58 symptomatic AEs which were not systematically administered to all patients, a new survey containing exactly these 58 symptomatic AEs was administered to all subsequently enrolled patients.

## Procedure

PRO-CTCAE items were completed by participants prior to clinic appointments on tablet computers via the PRO-CTCAE measurement system hosted on a secure server at the NCI.<sup>17</sup> To optimize usability by individuals with disabilities, PRO-CTCAE software is compliant with Section 508 of the U.S. Rehabilitation Act. The PRO-CTCAE measurement system employs conditional branching for AEs that contain more than a single attribute, such that subsequent items about severity or interference are skipped if respondents indicate that they are not experiencing a specific symptomatic AE. Participants were required to answer questions without assistance, but could request technical assistance with using the tablet computer from study staff.

## Anchors

Anchors are measurable criteria pre-specified as comparators in an instrument validation study. Examples of anchors relevant in PRO validation studies include well-validated patient- and clinician-reported outcomes and clinical variables such as disease site or concurrent medication use. For this study, anchors selected a priori included both generic measures (e.g., patient-reported global health-related quality of life [HRQOL] or clinician-reported performance status) and more specific clinical variables (e.g., antiemetic use or receipt of taxane chemotherapy). These anchors were selected based on literature review, expert consensus, and patient representative input.

The PRO anchors were administered to participants using a paper booklet containing the European Organisation for Research and Treatment of Cancer Core Quality of Life Questionnaire (EORTC QLQ-C30),<sup>20</sup> a 30-item instrument which produces a HRQOL summary score,<sup>21,22</sup> a global health status/quality of life (QOL) scale score, 5 functioning (physical, role, emotional, social, cognitive) scale scores, and 9 selected symptom item/scale scores. 28 items are measured on a 1-4 scale (1=not at all; 4=very much) with the remaining two items (overall health and QOL) scored on a 1-7 scale (1=very poor; 7=excellent). Like PRO-CTCAE, the recall period for the QLQ-C30 is “the past week”. Patients also completed three Global Impression of Change (GIC)<sup>23,24</sup> items at the primary follow-up visit. These items asked patients to rate their changes in overall QOL; physical condition; and emotional state on a 7-point scale ranging from “very much better”, “moderately better”, “a little better”, “about the same”, “a little worse”, “moderately worse”, to “very much worse”.

Clinician-reported ECOG PS was collected at each visit via a case report form. Other clinical anchors were abstracted from medical charts and included whether the participant had received radiation, surgery, and/or chemotherapy in the prior two weeks; type of chemotherapy; and use of specific medication classes, including: hormonal therapy, narcotic

analgesics, laxatives/stool softeners, antiemetics, sleep aids, anti-diarrhea medications, antacids, bronchodilators/inhaled corticosteroids, anxiolytics, and/or antidepressants.

### Study Visits

Participants were assigned to one of three groups with differing questionnaire schedules based on cancer type and clinic visit schedule, to avoid the necessity of extra clinic visits in this symptomatic population (eFigure 1). Group A included patients undergoing daily radiation or chemoradiation to enable analyses of test-retest reliability and varying recall periods (recall period analyses will be reported separately).<sup>25</sup> Group B included patients with at least four planned consecutive weekly clinic visits. Group C included participants whose planned clinic visits precluded participation in Group B but who did have a return clinic visit planned within 1-6 weeks. Irrespective of group assignment, all patients completed PRO-CTCAE items and QLQ-C30 at two visits that were spaced approximately 1-6 weeks apart. At each visit, ECOG PS and other clinical anchors were recorded on case report forms. PRO-CTCAE surveys administered to patients in Group A on the business day following study day 1 were used for the analysis of test-retest reliability, and included 49 pre-specified PRO-CTCAE items.

### Statistical analysis

Construct validity reflects the association between a new measurement tool and an established measure of the underlying concept(s) of interest. Construct validity is often investigated through convergent validity, which determines if the new measure moves in the same direction as an established instrument, and known-groups validity, which determines if the measurement tool can distinguish between groups of patients who are thought to be distinct with respect to the underlying concept being measured. To assess convergent validity, Pearson correlations were computed between each PRO-CTCAE item and QLQ-C30 HRQOL summary and other functioning/symptom scale scores. To aid interpretation, QLQ-C30 HRQOL summary and functioning/global scales were reverse scored such that higher scores represent inferior outcomes, matching the direction of PRO-CTCAE items. Pearson correlation values of .1, .3, and .5 were interpreted as small, medium, and large.<sup>26</sup> To assess known-groups validity, two-sample t-tests for ordinal 0-4 scales and chi-squared tests for binary scales were used to compare each PRO-CTCAE item between patients with high and low performance status (ECOG PS 0-1 versus 2-4). Additional known-groups analyses were pre-specified for PRO-CTCAE items that were expected to be higher in one group of patients versus another on the basis of cancer type, treatment, or other clinically relevant characteristic (e.g., pain in the abdomen in patients with gastrointestinal versus lung cancers). Effect sizes (computed as the difference between group means divided by the pooled standard deviation [Cohen's *d*], or difference between twice the arcsine of the square root of each sample proportion [Cohen's *h*]) of .2, .5, and .8 were interpreted as small, medium, and large.<sup>26</sup>

Test-retest reliability was estimated using the intra-class correlation coefficient (ICC) based on a one-way analysis of variance model<sup>27</sup> with an ICC of .7 or greater interpreted as acceptable.<sup>28</sup> Responsiveness of items was investigated by comparing change from first to second visit in 27 PRO-CTCAE items selected a priori. Comparisons were made using a

one-sided Jonckheere-Terpstra test across respondents who reported their GIC to be worse (“a little worse”, “moderately worse”, or “very much worse”), unchanged (“about the same”), or improved (“a little better”, “moderately better”, or “very much better”).<sup>29</sup> Standardized response means (SRM) were computed as the mean change score divided by the standard deviation of the change scores within each change category (worse versus no change versus improved) for each PRO-CTCAE item. Pearson correlations were also computed between PRO-CTCAE item changes and QLQ-C30 scale changes. One GIC item and one QLQ-C30 scale were specified a priori for each of the 27 PRO-CTCAE items. See eTable 2 for symptomatic AEs included in each analysis.

To accommodate conditional branching in the PRO-CTCAE software, values for automatically skipped items were assumed to be zero. P-values <.05 were considered statistically significant throughout. To take into consideration potential collinearity and multiplicity, sensitivity analyses employed a stricter p-value cut-off of <.001 and Hochberg's step-up procedure<sup>30</sup> across construct validity analyses within each item. An item was considered valid if statistical significance ( $P<.05$ ) along with a meaningful effect size (Pearson  $r .1$  or group difference effect size  $d$  or  $h .2$ ) was observed for at least one convergent or known-groups validity analysis.

## Results

Between January 2011 and February 2012, 975 patients initiating or undergoing chemotherapy and/or radiation were enrolled with 940/975 (96%) eligible patients completing PRO-CTCAE items at Visit 1 and 852/940 (91%) completing PRO-CTCAE items at Visit 2 (eFigure 1). Characteristics of the 940 participants included in this analysis are presented in Table 2. Median age was 59 years (range 19-91), 539 (57%) were female, 161 (17%) had impaired PS (ECOG 2-4), and 305 (32%) had no more than a high school education.

Most participants (938/940 [99.8%]) reported presence of at least one symptom (i.e., a score greater than 0) during the two primary visits, with 768/940 (82%) reporting at least one symptom as frequent, severe, and/or interfering “quite a bit” with daily activities. Patients were broadly symptomatic reporting presence of a median of 23 symptoms (range 0-91) with 904/940 (96%) reporting presence of 5 or more symptoms at the first visit. 118/124 (95%) PRO-CTCAE items were reported as present by at least 10% of respondents at both primary visits, with 82/124 (66%) items having at least 25% prevalence. The distribution of item scores for the set of 20 “core” symptomatic AEs appears in eFigure 2.

Detailed results related to construct validity of PRO-CTCAE items using all anchors are provided in eTable 3. With respect to convergent validity, 122/124 (98%) PRO-CTCAE items were associated in the expected direction with the QLQ-C30 HRQOL summary score (102/124  $P<.05$ ; 87/124  $P<.001$ ; Figure 1); 107/124 items demonstrated meaningful correlation (Pearson  $r .1$ ). When considering all QLQ-C30 functioning/global scales, 124/124 (100%) PRO-CTCAE items were associated in the expected direction with one or more scales, with 114/124 demonstrating meaningful correlation (Pearson  $r .1$ ), and 111/124 coefficients reaching statistical significance ( $P<.05$ ; 90/124  $P<.001$ ). PRO-CTCAE

items that were likely to impact physical functioning had the strongest correlations with the QLQ-C30 physical functioning scale (e.g., shortness of breath severity: Pearson  $r=.47$ ,  $P<.001$ ) whereas items likely to impact cognitive functioning had the strongest correlations with the QLQ-C30 cognitive functioning scale (e.g., problems with concentration severity: Pearson  $r=.71$ ,  $P<.001$ ; problems with memory severity: Pearson  $r=.69$ ,  $P<.001$ ). Similar results were seen between PRO-CTCAE items and conceptually-related QLQ-C30 emotional, role, and social functioning scales. For those PRO-CTCAE items with a parallel QLQ-C30 symptom scale/item (e.g., fatigue), large correlations between analogous items (all Pearson  $r>.69$ ,  $P<.001$ ) were consistently observed.

In the known-groups comparison between patients with low and high performance status, 94/124 PRO-CTCAE items had higher mean scores in the ECOG PS 2-4 group versus 0-1 group (58/124,  $P<.05$ ; 37/124,  $P<.001$ ; shown for 37 PRO-CTCAE items in eFigure 3).

In 127 a priori known-groups comparisons involving 87 PRO-CTCAE items based on cancer type, treatment, or other clinically relevant characteristic, 110/127 comparisons demonstrated higher PRO-CTCAE scores in the group expected to have worse symptom experience (85/127,  $P<.05$ ; 53/127,  $P<.001$ , eTable 3).

Most PRO-CTCAE items (119/124) reached a statistically significant and meaningful effect size on one or more construct validity criteria. The five items that did not exhibit at least one statistically significant and meaningful effect had low prevalence in this sample, thereby limiting our analysis. These items were: nosebleeds (prevalence 14.9% [frequency] and 14.0% [severity]); pain, swelling or redness at site of drug injection or intravenous therapy (prevalence 12.5%); pain during vaginal sex (prevalence 20.7%); and rash (prevalence 17.5%). A majority of PRO-CTCAE items (99/124 and 101/124) remained statistically significant under stricter criteria ( $P<.001$  and Hochberg's  $P<.05$ ) in sensitivity analyses (eTable 3).

In the subset of 80 respondents who completed PRO-CTCAE on consecutive business days (median 1 day, range 1-3 days), the test-retest reliability for the 49 pre-specified items ranged from .53 to .96 (median ICC .76) with 36/49 items having an ICC  $>.7$  (eTable 4).

In the analysis of responsiveness (Figure 2), statistically significant ( $P<.05$ ) monotonically decreasing mean PRO-CTCAE change scores were observed for 23 of 27 pre-specified items ( $P<.001$  for 13 items). The median SRM in patients reporting worsening was .19 (range .03-.40), whereas median SRM in patients reporting improvement was -.14 (range -.30-.09). Statistically significant correlations were observed between PRO-CTCAE item changes and corresponding QLQ-C30 scale changes for all 27 pre-specified items (median  $r=.43$ , range .10-.56; all  $P <.006$ ).

## Discussion

This large-scale multicenter study in adults undergoing active cancer therapy provides evidence supporting the validity, reliability, and responsiveness of the items in the PRO-CTCAE library. The PRO-CTCAE is unique in its intended use to complement the CTCAE

by providing comprehensive data on symptomatic AEs in cancer clinical trials from the patient perspective.

The design of this study posed a unique methodological challenge, due to the goal of assessing, within a single investigation, the measurement properties of 124 individual items representing a broad spectrum of symptomatic toxicities. Typically, PRO validation studies will test the properties of a single composite index score or a small number of domains that encompass related concepts. For the assessment of validity in the current study, the primary strategy to address this challenge was inclusion of both broad generic anchors (e.g., global HRQOL, ECOG PS) and more specific clinical variables (e.g., receipt of specific medication classes such as antiemetics). Interestingly, all of the PRO-CTCAE items were associated in the expected direction with at least one generic functioning measure, suggesting the impact that even a single toxicity may have on the patient experience.

Strengths of this study include the diverse sample, reflecting a wide range of cancer types and treatment modalities, and enrichment for less common cancer types. The sample was also successfully enriched for patients with impaired performance status (ECOG PS 2), enabling demonstration of the meaningfulness of PRO-CTCAE among those with substantial symptom burdens, as well as the feasibility of survey administration in debilitated patients. Moreover, participants were accrued at both academic and community sites across the U.S., including rural and urban settings, and reflected a range of educational and racial backgrounds.

Several caveats should be considered. First, our study was conducted in an English-speaking U.S.-residing patient population. Ongoing research is evaluating linguistic adaptations of PRO-CTCAE, and the measurement properties of both the English and other language versions in settings outside the U.S.<sup>31</sup> Linguistic validation of a Spanish language translation of PRO-CTCAE is being reported elsewhere.<sup>18</sup> Second, we assessed reliability in a subset of 49 items; thus, future studies to examine the test-retest reliability of the remaining PRO-CTCAE items are warranted. Third, a small number of highly specific symptomatic AEs were uncommon in the study sample and received low endorsement rates, thus limiting our ability to evaluate their measurement properties. Specifically, five items reflecting four symptomatic AEs (nosebleeds; pain/swelling/redness at site of drug injection or intravenous therapy; pain during vaginal sex; and rash) did not exhibit a statistically significant and meaningful effect on at least one construct validity criterion. These items are being evaluated in other clinical trial contexts. While the large number of items and anchors evaluated in this study raises the possibility of inflated Type I error, in sensitivity analyses using more stringent significance thresholds, the majority of items retained statistical significance. Lastly, notwithstanding inclusion of diverse malignancies in this study, results may not fully generalize to populations with rare tumor types. However, a prior cognitive interviewing study<sup>16</sup> affirms that PRO-CTCAE items were well understood by respondents with varying disease sites and receiving diverse anti-cancer treatments. Continued evaluation of PRO-CTCAE is currently underway in a variety of trial contexts to support the interpretability and value of patient-reporting of symptomatic treatment-related toxicity.

The CTCAE has historically enabled clinicians to describe the toxicity burden of cancer treatments using a consistent standard language allowing comparisons across trials. The value of patients' input in describing their own experiences is well recognized. Having a measurement system which integrates the patient perspective into AE reporting and which fosters consistency, transparency, and comparability across trials is similarly an important objective. The results of this validation study suggest that PRO-CTCAE can achieve its intended aim of integrating the patient experience into routine clinical trial AE reporting thereby augmenting the capacity for informed decision-making. In conclusion, this large-scale multicenter validation study in individuals undergoing active cancer therapy provides robust evidence for the validity, reliability, and responsiveness of items in the PRO-CTCAE library.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, and approval of the manuscript was supported by contracts from the U.S. National Cancer Institute: HHSN261200800043C, HHSN261201000063C, and HHSN261200800001E. The authors are grateful to the patients and clinicians who participated in this study. We thank Gitana Davila and Kathy Alexander for their contributions to the study including study coordination, patient recruitment, and data management at The Center for Cancer Prevention and Treatment at St. Joseph Hospital of Orange (Orange, CA, USA) and Hartford Hospital-Helen and Harry Gray Cancer Center (Hartford, CT, USA), respectively. We also thank Dr. Neil Aaronson for his methodological input on the EORTC QLQ-C30. The authors declare that they have no competing interests in relation to this research. Dr. Dueck had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

## References

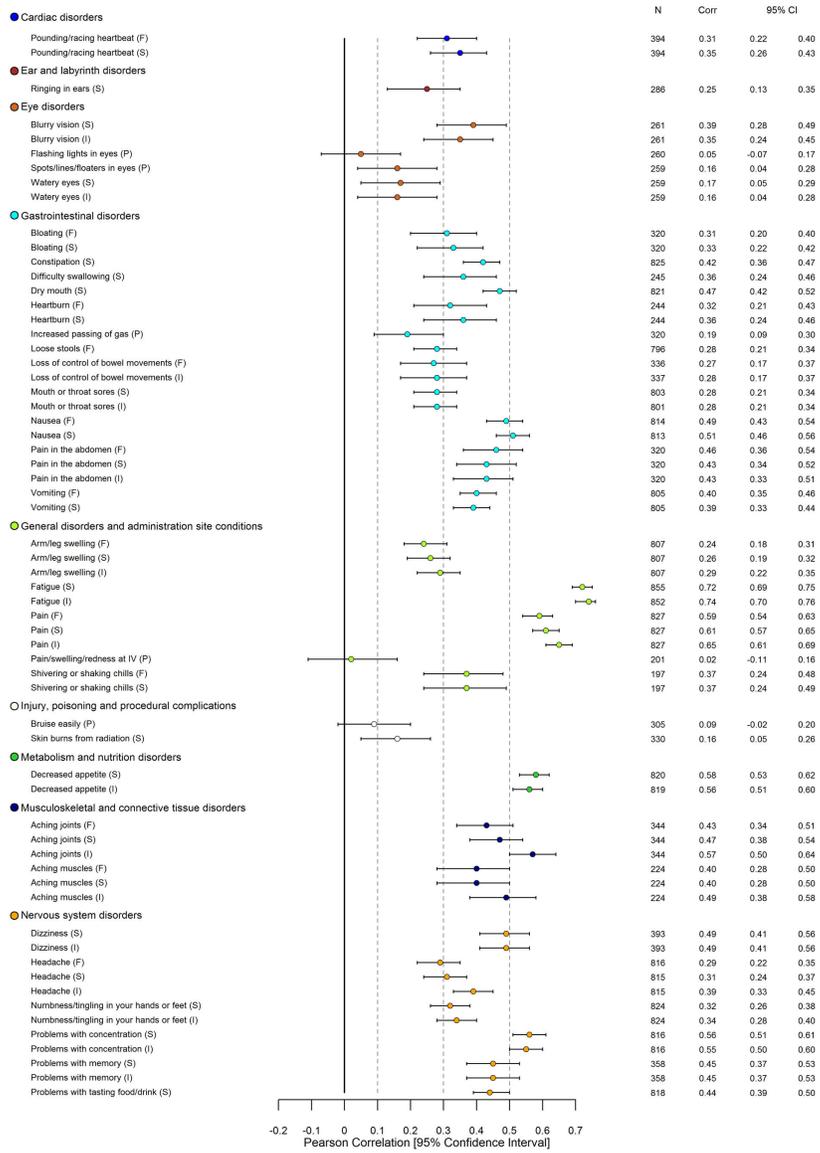
1. National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services. [last accessed March 16, 2015] Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0. NIH publication # 09-7473. Published May 29, 2009; Revised Version 4.03 June 14, 2010. Available at [http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE\\_4.03\\_2010-06-14\\_QuickReference\\_5x7.pdf](http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_5x7.pdf)
2. National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services. [last accessed March 16, 2015] caBIG Knowledge Center, CTCAE FAQ. Available at [http://evs.nci.nih.gov/ftp1/CTCAE/Archive/CTCAE\\_4.01\\_2009-07-14\\_FAQ.doc](http://evs.nci.nih.gov/ftp1/CTCAE/Archive/CTCAE_4.01_2009-07-14_FAQ.doc)
3. Trotti A, Colevas AD, Setser A, Basch E. Patient-Reported Outcomes and the Evolution of Adverse Event Reporting in Oncology. *J Clin Oncol.* 2007; 25(32):5121–7. [PubMed: 17991931]
4. Atkinson TM, Li Y, Coffey CW, Sit L, Shaw M, Lavene D, Bennett AV, Fruscione M, Rogak L, Hay J, Gönen M, Schrag D, Basch E. Reliability of adverse symptom event reporting by clinicians. *Qual Life Res.* 2012; 21:1159–1164. [PubMed: 21984468]
5. Basch E. The Missing Voice of Patients in Drug-Safety Reporting. *N Engl J Med.* 2010; 362(10): 865–869. [PubMed: 20220181]
6. Fromme EK, Eilers KM, Mori M, Hsieh YC, Beer TM. How accurate is clinician reporting of chemotherapy adverse effects? A comparison with patient-reported symptoms from the Quality-of-Life Questionnaire C30. *J Clin Oncol.* 2004; 22(17):3485–90. [PubMed: 15337796]
7. Pakhomov SV, Jacobsen SJ, Chute CG, Roger VL. Agreement between patient-reported symptoms and their documentation in the medical record. *Am J Manag Care.* 2008; 14(8):530–9. [PubMed: 18690769]

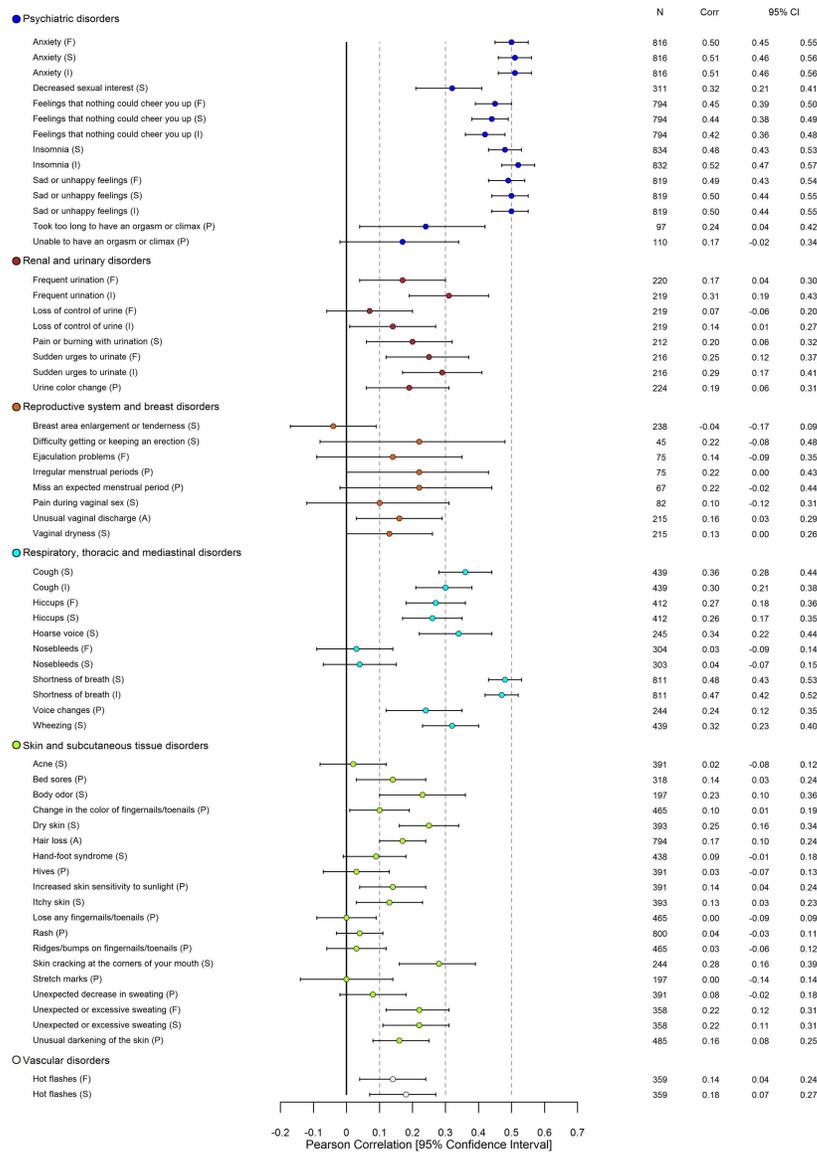
8. Basch E, Jia X, Heller G, Barz A, Sit L, Fruscione M, Appawu M, Iasonos A, Atkinson T, Goldfarb S, Culkin A, Kris MG, Schrag D. Adverse symptom reporting by patients versus clinicians: relationships with clinical outcomes. *J Nat Cancer Inst.* 2009; 101(23):1–9.
9. Quinten C, Maringwa J, Gotay CC, Martinelli F, Coens C, Reeve BB, Flechtner H, Greimel E, King M, Osoba D, Cleeland C, Ringash J, Schmucker-Von Koch J, Taphoorn MJ, Weis J, Bottomley A. Patient self-reports of symptoms and clinician ratings as predictors of overall cancer survival. *J Natl Cancer Inst.* 2011; 103(24):1851–8. [PubMed: 22157640]
10. Meacham, R.; McEngart, D.; O'Gorman, H.; Wenzel, K. Use and compliance with electronic patient reported outcomes within clinical drug trials; Presented at 2008 ISPOR (International Society for Pharmacoeconomics and Outcomes Research) Annual International Meeting; Toronto, Ontario, Canada.
11. Judson TJ, Bennett AV, Rogak LJ, Sit L, Barz A, Kris MG, Hudis CA, Scher HI, Sabbatini P, Schrag D, Basch E. Feasibility of long-term patient self-reporting of toxicities from home via the Internet during routine chemotherapy. *J Clin Oncol.* 2013 Jul 10; 31(20):2580–5. [PubMed: 23733753]
12. Di Maio M, Gallo C, Leigh NB, Piccirillo MC, Daniele G, Nuzzo F, Gridelli C, Gebbia V, Ciardiello F, De Placido S, Ceribelli A, Favaretto AG, de Matteis A, Feld R, Butts C, Bryce J, Signoriello S, Morabito A, Rocco G, Perrone F. Symptomatic toxicities experienced during anticancer treatment: agreement between patient and physician reporting in three randomized trials. *J Clin Oncol.* 2015; 33(8):910–5. [PubMed: 25624439]
13. Basch E, Iasonos A, McDonough T, Barz A, Culkin A, Kris MG, Scher HI, Schrag D. Patient versus clinician symptom reporting using the National Cancer Institute Common Terminology Criteria for Adverse Events: results of a questionnaire-based study. *Lancet Oncol.* 2006; 7(11): 903–9. [PubMed: 17081915]
14. Basch EM, Reeve BB, Mitchell SA, Clauser SB, Minasian L, Sit L, Chilukuri R, Baumgartner P, Rogak L, Blauel E, Abernethy AP, Bruner D. Electronic toxicity monitoring and patient-reported outcomes. *Cancer J.* 2011 Jul-Aug;17(4):231–4. [PubMed: 21799330]
15. Basch E, Reeve BB, Mitchell SA, Clauser SB, Minasian LM, Dueck AC, Mendoza TR, Hay J, Atkinson TM, Abernethy AP, Bruner DW, Cleeland CS, Sloan JA, Chilukuri R, Baumgartner P, Denicoff A, St Germain D, O'Mara AM, Chen A, Kelaghan J, Bennett AV, Sit L, Rogak L, Barz A, Paul DB, Schrag D. Development of the National Cancer Institute 's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *J Natl Cancer Inst.* 2014 Sep 29.106(9)
16. Hay JL, Atkinson TM, Reeve BB, Mitchell SA, Mendoza TR, Willis G, Minasian LM, Clauser SB, Denicoff A, O'Mara A, Chen A, Bennett AV, Paul DB, Gagne J, Rogak L, Sit L, Viswanath V, Schrag D, Basch E. NCI PRO-CTCAE Study Group. Cognitive interviewing of the U.S. National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). *Qual Life Res.* 2014 Feb; 23(1):257–69. [PubMed: 23868457]
17. Agarwal, H.; Gulati, M.; Baumgartner, P.; Coffey, D.; Kumar, V.; Chilukuri, R.; Shouery, M.; Reeve, B.; Basch, E. Web interface for the patient-reported version of the CTCAE; NCI caBIG Annual Meeting; July 2009; Washington, DC.
18. Arnold, BJ.; Mitchell, SA.; Lent, L.; Mendoza, TR.; Rogak, L.; Barragan, N.; Willis, GB.; Medina, M.; Lechner, SC.; Harness, J.; Basch, E. Linguistic validation of the Spanish translation of the U S National Cancer Institute's Patient- Reported Outcomes version of the Common Terminology Criteria for Adverse Events; Presentation. International Society for Quality of Life Research 21st Annual Conference; Berlin, Germany. October 15-18, 2014;
19. Reeve BB, Mitchell SA, Dueck AC, Basch E, Cella D, Reilly CM, Minasian LM, Denicoff AM, O'Mara AM, Fisch MJ, Chauhan C, Aaronson NK, Coens C, Bruner DW. Recommended patient-reported core set of symptoms to measure in adult cancer treatment trials. *J Natl Cancer Inst.* 2014 Jul 8.106(7)
20. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, Filiberti A, Flechtner H, Fleishman SB, de Haes JC, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst.* 1993 Mar 3; 85(5):365–76. [PubMed: 8433390]

21. Gundy CM, Fayers PM, Groenvold M, Petersen MA, Scott NW, Sprangers MA, Velikova G, Aaronson NK. Comparing higher order models for the EORTC QLQ-C30. *Qual Life Res.* 2012 Nov; 21(9):1607–17. [PubMed: 22187352]
22. Kieffer, JM.; Giesinger, JM.; Fayers, P.; Groenvold, M.; Peterson, MA.; Scott, NW.; Sprangers, MAG.; Velikova, G.; Aaronson, NK. Replication and relative validity of the mental and physical higher order model of the EORTC QLQ-C30; Presentation. International Society for Quality of Life Research 21st Annual Conference; Berlin, Germany. October 15-18, 2014;
23. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol.* 1998 Jan; 16(1):139–44. [PubMed: 9440735]
24. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Clinical Significance Consensus Meeting Group. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc.* 2002 Apr; 77(4):371–83. [PubMed: 11936935]
25. Mendoza, TR.; Bennett, AV.; Mitchell, SA.; Reeve, BB.; Atkinson, TM.; Li, Y.; Rogak, L.; Dueck, AC.; Basch, E. Impact of recall period on the accuracy of selected items from the US National Cancer Institute's Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE); Poster. International Society for Quality of Life Research 21st Annual Conference; Berlin, Germany. October 15-18, 2014;
26. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.
27. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin.* 1979; 86:420–8. [PubMed: 18839484]
28. Nunnally, JC.; Bernstein, IH. *Psychometric Theory.* 3rd. New York: McGraw-Hill; 1994.
29. Jonckheere AR. A distribution-free k-sample test against ordered alternatives. *Biometrika.* 1954; 41:133–45.
30. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika.* 1988; 75:800–2.
31. Kirsch M, Mitchell SA, Dobbels F, Stussi G, Basch E, Halter JP, De Geest S. Linguistic and content validation of a German-language PRO-CTCAE-based patient-reported outcomes instrument to evaluate the late effect symptom experience after allogeneic hematopoietic stem cell transplantation. *Eur J Oncol Nurs.* 2015 Feb; 19(1):66–74. [PubMed: 25190633]

### At A Glance

- Symptomatic adverse events (AEs) in cancer trials are currently graded by clinicians using the National Cancer Institute's (NCI) Common Terminology Criteria for Adverse Events (CTCAE)
- This study assessed the measurement properties (validity, reliability, and responsiveness) of the newly developed NCI Patient-Reported Outcomes version of the CTCAE (PRO-CTCAE)
- A large, heterogeneous sample of 940 adult cancer patients undergoing outpatient cancer treatment provided PRO-CTCAE and other patient-reported and clinical data
- A majority of the PRO-CTCAE items (119 out of 124) met at least a validity criterion
- PRO-CTCAE provides a valid and reliable assessment of symptomatic toxicities from the patient's perspective, and is encouraged for use in oncology trials to enhance the accuracy of AE reporting

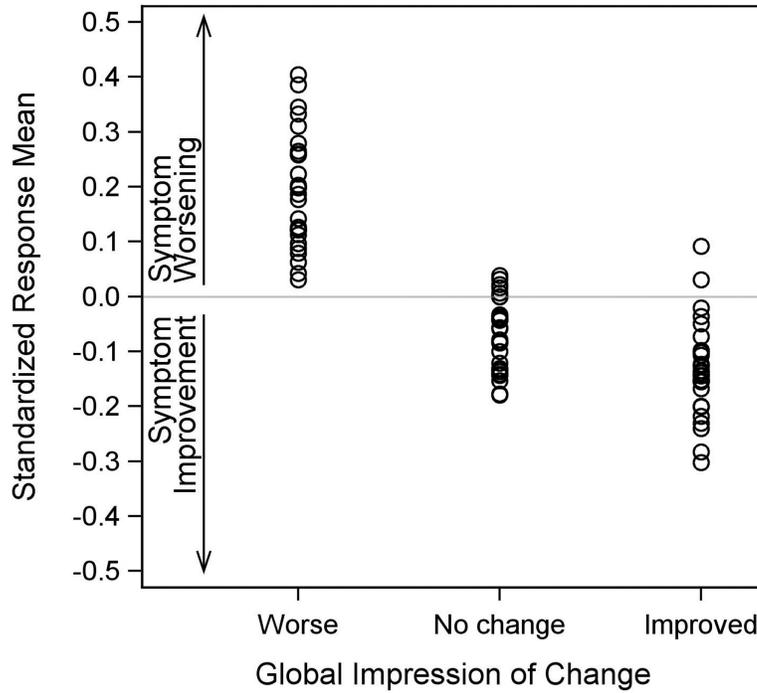




**Figure 1. Pearson Correlations between 124 PRO-CTCAE Item Scores and EORTC QLQ-C30 HRQOL Summary Score\* at Visit 1**

Abbreviations: PRO-CTCAE, Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events; EORTC QLQ-C30, European Organisation for Research and Treatment of Cancer Core Quality of Life Questionnaire; HRQOL, Health-related quality of life; CTCAE, Common Terminology Criteria for Adverse Events

\*See eTable 3 for all computed Pearson correlations between PRO-CTCAE items and EORTC QLQ-C30 functioning, global, and symptom scales.



**Figure 2. Standardized Response Means across 27 PRO-CTCAE Items by Patient-Reported Global Impression of Change Category\***

Abbreviation: PRO-CTCAE, Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events

\*Figure 2 includes 27 frequency, severity, and interference items selected prior to initiation of the responsiveness analysis. The set of 20 “core” symptomatic AEs was reviewed and symptomatic AEs were selected if they had high potential to be meaningfully related to global changes in quality of life, physical condition, and/or emotional state (i.e., the Global Impression of Change items which were administered at the second visit). Of the 20 reviewed symptomatic AEs, 13 were included based on this criterion (see eTable 2). The symptomatic AEs which were excluded were felt to be related to initiation or changes in specific treatments (dry mouth, problems with tasting food/drink, rash) so may not exhibit change in a heterogeneously treated sample of patients; require a longer duration of follow-up to exhibit change (arm/leg swelling, hair loss); or be related to cognitive condition (headache, problems with concentration) which was not assessed in the Global Impression of Change items.

**Table 1**  
**PRO-CTCAE Item Formats\***

Please think back over the past 7 days:	Example
<b>Frequency</b> (25 symptomatic AE terms): How <b>OFTEN</b> did you have _____? Never / Rarely / Occasionally / Frequently / Almost constantly	Vomiting
<b>Severity</b> (51 symptomatic AE terms): What was the <b>SEVERITY</b> of your _____ at its WORST? None / Mild / Moderate / Severe / Very severe	Pain
<b>Interference</b> (25 symptomatic AE terms): How much did _____ <b>INTERFERE</b> with your usual or daily activities? Not at all / A little bit / Somewhat / Quite a bit / Very much	Sudden urges to urinate
<b>Presence</b> (21 symptomatic AE terms): Did you have any _____? No / Yes	Unusual darkening of the skin
<b>Amount</b> (2 symptomatic AE terms): Did you have any _____? Not at all / A little bit / Somewhat / Quite a bit / Very much	Hair loss

Abbreviations: PRO-CTCAE, Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events; AE, adverse event

\* See Basch et al.<sup>15</sup> for a complete listing of PRO-CTCAE items.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**  
**Patient Characteristics (N=940)**

Characteristic	No.	%
Age at enrollment		
Median	59	
Range	19 - 91	
Age group		
<30	23	2.5%
30-64	597	63.5%
65-74	235	25.0%
75	85	9.0%
Gender		
Female	539	57.3%
Male	401	42.7%
Ethnicity		
Hispanic or Latino	56	6.0%
Not Hispanic or Latino	832	88.5%
Missing	52	5.5%
Race		
White	675	71.8%
Black or African American	203	21.6%
Asian	42	4.5%
Other or multiple races reported	8	0.9%
Missing	12	1.3%
Education		
High school or less	305	32.4%
Some college	199	21.2%
College graduate or more	415	44.1%
Missing	21	2.2%
Cancer type		
Lung, head or neck	329	35.0%
Breast	260	27.7%
Genitourinary or gynecologic	172	18.3%
Gastrointestinal	95	10.1%
Hematologic	47	5.0%
Other or unknown	37	3.9%
ECOG Performance Status at first visit		
0-1	779	82.9%
2-4	161	17.1%

Characteristic	No.	%
Cancer treatment in prior two weeks		
Chemotherapy	522	55.5%
Radiation	424	45.1%
Surgery	35	3.7%

Abbreviation: ECOG, Eastern Cooperative Oncology Group

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript