



Published in final edited form as:

Clin Trials. 2017 June ; 14(3): 255–263. doi:10.1177/1740774517698645.

Evaluation of different recall periods for the US National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE)

Tito R. Mendoza, PhD¹, Amylou C. Dueck, PhD², Antonia V. Bennett, PhD³, Sandra A. Mitchell, PhD⁴, Bryce B. Reeve, PhD³, Thomas M. Atkinson, PhD⁵, Yuelin Li, PhD⁵, Kathleen M. Castro, RN, MSN⁴, Andrea Denicoff, RN, MS⁴, Lauren J. Rogak, MA⁵, Richard L. Piekarz, MD, PhD⁴, Charles S. Cleeland, PhD¹, Jeff Sloan, PhD⁶, Deborah Schrag, MD⁷, and Ethan Basch, MD, MSc^{3,5}

¹University of Texas, MD Anderson Cancer Center, Houston, TX, USA

²Mayo Clinic, Scottsdale, AZ, USA

³Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, USA

⁴National Cancer Institute, Rockville, MD, USA

⁵Memorial Sloan Kettering Cancer Center, New York, NY, USA

⁶Mayo Clinic, Rochester, MN, USA

⁷Dana-Farber Cancer Institute, Boston, MA, USA

Abstract

Aims—The U.S. National Cancer Institute recently developed the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). PRO-CTCAE is a library of questions for clinical trial participants to self-report symptomatic adverse events (e.g., nausea). The objective of this study is to inform evidence-based selection of a recall period when PRO-CTCAE is included in a trial. We evaluated differences between 1-week, 2-week, 3-week, and 4-week recall periods, using daily reporting as the reference.

Methods—English-speaking patients with cancer receiving chemotherapy and/or radiotherapy were enrolled at four U.S. cancer centers and affiliated community clinics. Participants completed 27 PRO-CTCAE items electronically daily for 28 days, and then weekly over 4 weeks, using 1-week, 2-week, 3-week, and 4-week recall periods. For each recall period, mean differences, effect sizes, and intraclass correlation coefficients were calculated to evaluate agreement between the maximum of daily ratings and the corresponding ratings obtained using longer recall periods (e.g., maximum of daily scores over 7 days vs. 1-week recall). Analyses were repeated using the average of daily scores within each recall period rather than the maximum of daily scores.

Corresponding author: Tito Mendoza, PhD, Department of Symptom Research, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 1450, Houston, TX 77030, tmendoza@mdanderson.org, Tel: 713-745-3486, Fax: 713-745-3475.

Conflict of Interest

The Authors declare that there are no conflicts of interest.

Results—127 subjects completed questionnaires (57% male; median age 57). The median of the 27 mean differences in scores on the PRO-CTCAE 5-point response scale comparing the maximum daily versus the longer recall period (and corresponding effect size), was -0.20 (-0.20) for 1-week recall; -0.36 (-0.31) for 2-week recall; -0.45 (-0.39) for 3-week recall; and -0.47 (-0.40) for 4-week recall. The median intraclass correlation across 27 items between the maximum of daily ratings and the corresponding longer recall ratings for 1-week recall was 0.70 (range: 0.54–0.82); 2-week recall: 0.74 (range: 0.58–0.83); 3-week recall: 0.72 (range: 0.61–0.84); and 4-week recall: 0.72 (range: 0.64–0.86). Similar results were observed for all analyses using the average of daily scores rather than the maximum of daily scores.

Conclusions—1-week recall corresponds best to daily reporting. Although intraclass correlations remain stable over time, there are small but progressively larger differences between daily and longer recall periods at 2, 3, and 4 weeks, respectively. The preferred recall period for the PRO-CTCAE is the past 7 days, although investigators may opt for recall periods of 2, 3, or 4 weeks with an understanding that there may be some information loss.

Trial registration—ClinicalTrials.gov NCT02158637

Keywords

Recall Period; Patient-Reported Outcomes; Symptomatic Adverse Events; Validation; Measurement Properties; PRO-CTCAE

Introduction

The U.S. National Cancer Institute Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) is a recently developed library of self-report items representing 78 symptomatic toxicities designed to capture symptomatic adverse events in cancer clinical trials.¹ When the PRO-CTCAE was developed, a default recall period of the “past 7 days” was selected.² This decision was based on both practical and theoretical grounds; specifically that daily assessment would not be feasible in most cancer clinical trials, and concern that recall periods longer than one week might lead to meaningful degradation of memory and resultant loss of information.³ Nonetheless, it was recognized that longer recall periods might be desired or necessary for logistical reasons in some clinical trials, for example when PRO-CTCAE questionnaires are administered to participants only at clinic visits which are spaced out longer than 7 days during active therapy.

The choice of recall period is recognized as an important consideration when evaluating the measurement properties of patient-reported outcome questionnaires. For example, the U.S. Food and Drug Administration’s guidance on the use of patient-reported outcome measures to support product labeling claims recommends that selection of a given recall period should be justified, and that shorter recall periods are generally preferred.⁴ The recall period for a measure can be determined based on the outcomes of interest, anticipated characteristics of study participants, and logistical issues.⁵ Although self-reports are influenced by cognitive heuristics,⁶ prior research has found that 1-day recall is comparable to the average momentary assessment for the day.³ Studies of patient-reported outcome questionnaires in

cystic fibrosis,⁷ type 2 diabetes,⁸ and chronic obstructive pulmonary disease⁹ found 7-day recall produced estimates that were somewhat larger than the average or worst of 7 daily reports. A study of pain and fatigue interference items¹⁰ observed that 3-day, 7-day, and 28-day recall were highly correlated but consistently larger than the average of daily reports from that period.

A recent study evaluated the concordance between daily and weekly symptom severity reports using a subset of PRO-CTCAE items in hematopoietic stem cell transplant recipients.¹¹ However the influence of longer recall periods on responses to PRO-CTCAE frequency, severity, and interference items in a diverse sample of patients undergoing cancer treatment has not been examined. The goal of the present study was to compare 1-, 2-, 3-, and 4-week recall to daily ratings of PRO-CTCAE symptom items in a large, multi-center study of cancer patients receiving or initiating chemotherapy, radiation therapy, or both.

Methods

Setting and sample

Adult patients with a solid tumor or hematologic malignancy receiving chemotherapy and/or radiation therapy at one of four U.S.-based cancer centers and their affiliated community clinics were eligible to participate in this study (Dana-Farber Cancer Institute, Boston, MA; Mayo Clinic, Rochester, MN, Memorial Sloan Kettering Cancer Center, New York, NY; University of Texas M. D. Anderson Cancer Center, Houston, TX). Eligible participants were approached in clinic waiting areas and invited to participate. All participants could read, write, and comprehend English and were without clinically significant cognitive impairment based on site investigator judgment. Institutional review board approval was obtained at all sites and at the U.S. National Cancer Institute, and all participants completed written informed consent. This recall study was an embedded substudy nested in a larger validation study of the PRO-CTCAE which has been previously reported (clinicaltrials.gov identifier NCT02158637).¹²

Measures

The PRO-CTCAE item library is comprised of 124 patient-reported outcome items representing 78 symptomatic adverse events, with each symptom term assessed relative to one or more attributes including frequency (F), severity (S), and/or interference with usual or daily activities (I). The structure of these items has been described previously (for more information and to register to use PRO-CTCAE in a study, see: <http://healthcaredelivery.cancer.gov/pro-ctcae>).¹ For this study, we included 27 PRO-CTCAE items representing 14 symptomatic adverse events that were identified by the National Cancer Institute as being prevalent and clinically impactful across cancer populations.¹³ It was thought to be infeasible to compare recall periods for all 124 items in a single study, and we assumed that the results for these 27 clinically-relevant symptom items (which included frequency, severity and interference items) could be generalized to the complete PRO-CTCAE item library.

The 27 items evaluated in this study included anxiety [attributes: F, S, I], sad or unhappy feelings [F, S, I], constipation [S], loose stools [F], anorexia [S, I], nausea [F, S], vomiting [F, S], dry mouth [S], mouth or throat sores [S, I], shortness of breath [S, I], numbness/tingling in hands and feet [S, I], pain [F, S, I], fatigue [S, I], and insomnia [S]. PRO-CTCAE items measure frequency, severity, or interference with daily activities using a 0–4 rating scale (i.e., *frequency*: (0) never, (1) rarely, (2) occasionally, (3) frequently, (4) almost constantly; *severity*: (0) none, (1) mild, (2) moderate, (3) severe, (4) very severe; and *interference with daily activities*: (0) not at all, (1) a little bit, (2) somewhat, (3) quite a bit, (4) very much).

Study Design

Participants completed PRO-CTCAE items using a 24-hour recall on a daily basis over a 4-week period (28 days) via an automated telephone interactive voice response system. Participants also completed the same PRO-CTCAE items weekly via the web using a 1-week (i.e., 7-day) recall period at 1 week following their baseline enrollment visit; a 2-week recall at 2 weeks, a 3-week recall at 3 weeks, and a 4-week recall at 4 weeks. Patients were instructed to rate their symptoms during ‘the past 7 days’ for the 1-week recall, ‘the past 14 days’ for the 2-week recall, ‘the past 21 days’ for 3-week recall and ‘the past 28 days’ for the 4-week recall. Similarity of scores between the interactive voice response system and web modes of PRO-CTCAE administration has been previously established, suggesting that if there is a bias effect due to mode of administration, it is minimal.¹⁴ Both the interactive voice response system and web-based surveys allowed for conditional branching where participants were not asked about how a symptom interfered with their daily activities if they did not report experiencing that symptom. Participants were trained to use the PRO-CTCAE web and interactive voice response system systems at the time of enrollment. Participants were instructed to answer questions without assistance from others, although they could request technical assistance from study staff. Demographic and clinical variables were gathered in a case report form.

Statistical analyses

Two different approaches were employed to compare daily with longer recall period ratings. First, the maximum of daily ratings was calculated for each recall period of interest and compared to the longer recall period ratings. This approach was used both because analyses of adverse events in clinical trials generally tabulate the worst magnitude of each event experienced by each patient (i.e., maximum rating) and because the phrasing of the PRO-CTCAE items asks about worst magnitude during the recall/reference period. Second, the average of daily ratings was calculated for each recall period of interest and compared to the longer recall period ratings. This approach is more typical for patient-reported outcome analyses focused on quality of life where average scores over time are reported. For each of these approaches, the difference between daily ratings for each recall period and the corresponding longer recall periods was averaged across participants. This was conducted for each of the 27 PRO-CTCAE items individually and summarized across the 27 items using the median of each reported statistic. Effect sizes were calculated for paired differences based on Dunlap et al.¹⁵ with the absolute value of the effect size interpreted based on Cohen (0–<0.2 as trivial; 0.2–<0.5 as small; 0.5–<0.8 as moderate; and 0.8 as

large)¹⁶. The strength of the associations between daily reports and the longer recall periods were evaluated using intraclass correlation coefficients using the intraclass correlation coefficients (3, 1) in the notation of Shrout and Fleiss¹⁷ in a two-way analysis of variance model. Intraclass correlation coefficient is a measure of agreement that takes into account multiple ratings of the same phenomenon by a single rater. Specifically, “intraclass correlation coefficient (3, 1)” refers to a correlation where each rater completes a pair of assessments (in this case, assessments made by daily recall and by 1, 2, 3 or 4 week recall), and these two assessments are the only assessments under consideration by a given analysis and the rater is considered as a fixed effect.¹⁷

For the primary analysis, participants had to have completed at least one daily report during each week of the respective assessment period as well as the longer recall report of interest (e.g., at least 1 of 7 reports during weeks 1, 2, 3, and 4 plus the 4-week recall items). Sensitivity analyses were conducted requiring at least 4 of 7 daily reports during each week, and separately requiring only a single daily report in any week during the study.

Results

Between January 2011 and February 2012, 127 patients receiving chemotherapy and/or radiotherapy for advanced cancer, and who were enrolled in the multi-center PRO-CTCAE validation study, participated in this recall sub-study (Table 1). The sample had a median age of 57 years (range 20–77), and a majority were Caucasian (83%). All had received radiation therapy, chemotherapy or both in the two weeks prior to study enrollment.

Table 2 shows item-level differences between the maximum of daily ratings and the longer recall period ratings for each recall period of interest (i.e., 1-week, 2-week, 3-week, and 4-week). In comparing daily vs. 1-week recall, the median of differences for the 27 items was –0.20 (effect size –0.20), on the 5-point PRO-CTCAE item response scale. At the item level, mean differences ranged from 0.00 to –0.33. For 1-week recall, the smallest differences were observed for the severity and interference of vomiting while the largest differences were observed for severity and interference of fatigue. For 2-week recall, the median of the 27 mean differences was –0.36 (range –0.09 to –0.66; overall effect size –0.31); for 3-week recall, the median of the 27 mean differences was –0.45 (range –0.25 to –0.72; overall effect size –0.39); and for 4-week recall, the median of the 27 mean differences was –0.47 (range –0.14 to –0.73; overall effect size –0.40). See eFigure 1. In a separate analysis comparing the average of daily ratings with the various longer recall periods, differences and effect sizes were slightly smaller in magnitude overall, with a similar trend of increasing values with each successively longer recall period (Table 3).

Results were similar in two different sensitivity analyses requiring at least 4 of 7 daily reports during each week (N=100), and requiring only a single daily report in any week (N=126) (results shown in Supplemental Appendix eTable 1, online only). For each of the 4 recall periods, the medians of the 27 effect sizes had a negative value when the maximum of daily ratings was compared with longer recall period ratings, but were positive when comparing the average of daily ratings. This indicates that the ratings provided with 1, 2, 3 and 4-week recall fell below the maximum daily rating, but above the average daily rating.

This observation suggests that recall periods of 1 week or more produce attenuated assessments of the symptom at its worst, when compared to the maximum daily rating, a pattern which is consistent across the 4 recall periods and across PRO-CTCAE symptoms.

Intraclass correlation coefficients computed using the maximum of daily ratings and the corresponding longer recall period ratings are shown in Table 4, and were similar across recall periods. In comparing daily vs. 1-week recall, the median intraclass correlation coefficient across the 27 items was 0.70 with a range from 0.54 to 0.82. For 2-week recall, the median intraclass correlation coefficient was 0.74 (range 0.58 to 0.83); for 3-week recall, the median intraclass correlation coefficient was 0.72 (range: 0.61 to 0.84) and for 4-week recall, the median intraclass correlation coefficient was 0.72 (range: 0.64 to 0.86). Similar trends in the intraclass correlation coefficients between the average of daily ratings (instead of maximum of daily ratings) and the corresponding recall periods were observed (Supplemental Appendix eTable 2, online only). For the average daily ratings, the median intraclass correlation coefficients across the 27 items were 0.74, 0.73, 0.76 and 0.78 for 1-week, 2-week, 3-week, and 4-week recall period, respectively. Compliance rates with daily and weekly reporting are shown in eTable 3.

Discussion

This study evaluated the relationships between 1, 2, 3, and 4-week recall with daily reporting for 27 PRO-CTCAE items in patients receiving systemic therapy for advanced cancers. Mean differences and effect sizes were smallest for 1-week recall. There were small but progressively greater mean differences and effect sizes with each incrementally longer recall period. Correlations were high and similar between daily reports with all of the longer recalled reports. These findings support use of 1-week recall as the preferred recall period for the PRO-CTCAE.

In trials that employ in-clinic, paper-and-pencil PRO-CTCAE reporting it may only be logistics that may limit the feasibility of data collection to every 3 or 4 weeks. Clinical trials are often performed in multicenter networks where electronic patient-reported outcome administration systems are not yet widely available, and patient-reported outcome questionnaires may need to be administered using paper forms at clinic visits. In this circumstance, a longer recall period (e.g., 3- or 4-week recall) may be considered, to assure comprehensive capture of symptomatic adverse events without temporal breaks, particularly with cyclic treatment regimens. This study provides information for investigators considering use of a longer recall period, demonstrating that there is some loss of information (i.e., underestimation of the true worst symptom experience) that must be balanced with trial logistics. While the longer recall period appears to underestimate the true maximum (i.e., worst) symptom experience, the longer recall period appears to overestimate the true average symptom experience. In the future, as electronic patient-reported outcome systems are more commonly available for use in cancer clinical trials, it is anticipated that a 1-week recall will become increasingly feasible in a majority of trial contexts.

Concerns about temporal breaks introduced by a 1-week recall may be less pertinent in therapeutic contexts where symptomatic toxicities are expected to be stable or changing only

subtly (for example, as with chronically administered oral therapies, or during long term follow-up after completion of active treatment when toxicities are expected to have stabilized). For example, in a trial of prolonged daily oral therapy, investigators may employ weekly reporting during the initial period of treatment, with the time between assessments lengthened after adverse effects are expected to have stabilized (for example, weekly reporting for the first two cycles, followed by monthly reporting during subsequent cycles, yet still maintaining the 1-week recall period). Similarly, during long-term follow up after treatment completion, when treatment effects are likely to be stable or changing only subtly, less frequent data collection allowing for temporal breaks is reasonable (e.g., evaluation every 6 months for up to 3 years, with the same 1-week recall period as used during active therapy).

Several caveats should be considered in interpreting these study results. Accrual was limited to four U.S.-based cancer centers and their affiliated community clinics. Although study participants had diverse levels of educational attainment and 18% were non-white, there were few Hispanic participants, and PRO-CTCAE items were tested only in English. Our sample was comprised predominantly of individuals with lung or head and neck cancers receiving chemoradiotherapy. These individuals are more likely to receive daily treatment, which accommodated the daily reporting study design and assured a high likelihood that symptoms would be prevalent and variable during the study period. Patients with other cancer sites were eligible to participate if their treatment schedule met the requirement of four to five consecutive weekly clinic visits, and comprised 26% of the sample. Study investigators felt it would be too burdensome to request daily recall of all 124 PRO-CTCAE items, and therefore a subset of 27 PRO-CTCAE items reflecting common toxicities across cancer clinical trials was selected for this study. Based on a prior study of mode equivalence,¹⁴ the differences in responses obtained by web or IVR based data collection are known to be so small that the differences observed in this study between daily and weekly assessments are not meaningfully impacted by mode of data collection. Finally, while it is possible that daily symptom reporting alters the recall of symptoms over longer periods of time, this potential artifact cannot be avoided in recall period studies as it is a necessary component of the study design¹⁸. Nonetheless, the consistency of results across symptoms, and the similarity of our standardized mean differences to those observed by other investigators,^{3,7,8,9,10} suggest the generalizability of our results to other PRO-CTCAE items which also measure toxicity in cancer clinical trials.

In the primary analysis, there were 26/126 (21%) participants with fewer than 4/7 self-reports during all weeks of study participation. However, multiple sensitivity analyses, including an analysis with only the 100 participants with 4/7 reports in all weeks, found similar results (shown in eTable 1).

There was some observed variability at the item level in the effects of the differing recall periods that should also be considered when using specific PRO-CTCAE items in a given trial. For example, the effect sizes of the differences between maximum daily and 4-week recall ratings were at least moderate in size for anxiety frequency, constipation severity, dyspnea frequency, fatigue severity and interference, insomnia severity, and loose stools

frequency. Finally, given potential differences in the amount of information loss across recall periods, recall period should be standardized across arms in multi-arm comparative trials.

Conclusion

1-week recall corresponds well to daily reporting. Although correlations remain stable over time, there are small but progressively larger differences between daily and longer recall periods at 2, 3, and 4 weeks, respectively. When employing PRO-CTCAE items in a clinical trial, the preferred recall period is the past 7 days. Longer recall periods of 2, 3, or 4 weeks may be selected by investigators as dictated by study design or logistics, with an understanding that a longer recall period may be associated with some loss of information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Support: This project was supported by contract HHSN261200800043C from the U.S. National Cancer Institute

References

1. 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)
2. 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 US 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]
3. Broderick JE, Schwartz JE, Vikingstad G, Pribbernow M, Grossman S, Stone AA. The accuracy of pain and fatigue items across different reporting periods. *Pain.* 2008 Sep 30; 139(1):146–57. [PubMed: 18455312]
4. U.S. Department of Health and Human Services, Food and Drug Administration. [last accessed 9/07/15] Guidance for industry: Patient-reported outcomes measures: Use in medical product development to support labeling claims. Dec. 2009 Available at <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf>
5. Stull DE, Leidy NK, Parasuraman B, Chassany O. Optimal recall periods for patient-reported outcomes: challenges and potential solutions *Curr Med Res Opin.* 2009 Apr; 25(4):929–42.
6. Gorin, AA., Stone, AA. Recall biases and cognitive errors in retrospective self-reports: a call for momentary assessments. In: Baum, A.Revenson, T., Singer, J., editors. *Handbook of health psychology.* Mahwah, NJ: Erlbaum; 2001. p. 405-13.
7. Bennett AV, Patrick DL, Lymp JF, Edwards TC, Goss CH. Comparison of 7-day and repeated 24-hour recall of symptoms of cystic fibrosis. *J Cyst Fibros.* 2010 Dec; 9(6):419–24. [PubMed: 20817624]
8. Bennett AV, Patrick DL, Bushnell DM, Chiou CF, Diehr P. Comparison of 7-day and repeated 24-h recall of type 2 diabetes. *Qual Life Res.* 2011 Jun; 20(5):769–77. DOI: 10.1007/s11136-010-9791-5 [PubMed: 21088912]

9. Bennett AV, Amtmann D, Diehr P, Patrick DL. Comparison of 7-day recall and daily diary reports of COPD symptoms and impacts. *Value Health*. 2012 May; 15(3):466–74. Epub 2012 Feb 9. DOI: 10.1016/j.jval.2011.12.005 [PubMed: 22583457]
10. Broderick JE, Schneider S, Scharz JE, Stone AA. Interference with activities due to pain and fatigue: accuracy of ratings across different reporting periods. *Quality of Life Research*. 2010; 19(8):1163–1170. [PubMed: 20535565]
11. Wood WA, Deal AM, Bennett AV, Mitchell SA, Abernethy AP, Basch E, Bailey C, Reeve BB. Comparison of seven-day and repeated 24-hour recall of symptoms in the first 100 days after hematopoietic cell transplantation. *J Pain Symptom Manage*. 2015 Mar; 49(3):513–20. [PubMed: 25128608]
12. Dueck AC, Mendoza TR, Mitchell SA, Reeve BB, Castro KM, Rogak LJ, Atkinson TM, Bennett AV, Denicoff AM, O'Mara AM, Li Y, Clauser SB, Bryant DM, Bearden JD 3rd, Gillis TA, Harness JK, Siegel RD, Paul DB, Cleeland CS, Schrag D, Sloan JA, Abernethy AP, Bruner DW, Minasian LM, Basch E. National Cancer Institute PRO-CTCAE Study Group. Validity and Reliability of the US National Cancer Institute's Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). *JAMA Oncol*. 2015 Nov; 1(8):1051–9. [PubMed: 26270597]
13. 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)
14. Bennett AV, Dueck AC, Mitchell SA, Mendoza TR, Reeve BB, Atkinson TM, Castro KM, Denicoff A, Rogak LJ, Harness JK, Bearden JD, Bryant D, Siegel RD, Schrag D, Basch E. National Cancer Institute PRO-CTCAE Study Group. Mode equivalence and acceptability of tablet computer-, interactive voice response system-, and paper-based administration of the U.S. National Cancer Institute's Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). *Health Qual Life Outcomes*. 2016 Feb 19.14(1):24. [PubMed: 26892667]
15. Dunlap WP, Cortina JM, Vaslow JB, Burke MJ. Meta-analysis of experiments with matched groups or repeated measures designs. *Psychol Methods*. 1996; 1:170–7.
16. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. 2. Routledge; New York, NY: 1988.
17. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*. 1979; 86:420–8. [PubMed: 18839484]
18. Norquist JM, Girman C, Fehnel S, DeMuro-Mercon C, Santanello N. Choice of recall period for patient-reported outcome (PRO) measures: criteria for consideration. *Qual Life Res*. 2012 Aug; 21(6):1013–20. DOI: 10.1007/s11136-011-0003-8 [PubMed: 21909804]

Table 1

Participant Characteristics (N=127)

Characteristic	No.	%
Age (years) at enrollment		
Median	57	
Range	20–77	
Age group (in years) *		
<30	6	5%
30–64	91	72%
65–74	29	23%
75	1	1%
Gender		
Female	55	43%
Male	72	57%
Race		
White	104	83%
Black or African American	14	11%
Asian	6	5%
Native Hawaiian/Other Pacific Islander	1	1%
Missing	2	--
Ethnicity		
Hispanic or Latino	7	6%
Not Hispanic or Latino	117	94%
Missing	3	--
Education		
High school or less	24	21%
Some college	20	17%
College graduate or more	73	62%
Missing	10	--
Cancer type *		
Lung, head or neck	94	74%
Breast	11	9%
Gastrointestinal	7	6%
Hematologic	8	6%
Other	7	6%
Cancer treatment received in prior two weeks **		
Chemotherapy	84	66%
Radiation	90	71%

Characteristic	No.	%
Surgery	6	5%

ECOG, Eastern Cooperative Oncology Group Performance Status

* Percentages sum to a greater than 100% due to rounding.

** Participants may have received more than one treatment modality in the prior two weeks

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Differences between the maximum of daily ratings and longer recalled ratings, by recall period, for each PRO-CTCAE item (0–4 scale).

	Maximum of 7 days vs. 1-week recall (N=126*)		Maximum of 14 days vs. 2-week recall (N=118)		Maximum of 21 days vs. 3-week recall (N=109)		Maximum of 28 days vs. 4-week recall (N=93)	
	Mean Difference	Effect size	Mean Difference	Effect size	Mean Difference	Effect size	Mean Difference	Effect size
Anxiety (F)	-0.20	-0.20	-0.50	-0.49	-0.48	-0.43	-0.55	-0.50
Anxiety (S)	-0.28	-0.30	-0.49	-0.52	-0.53	-0.54	-0.48	-0.47
Anxiety (I)	-0.28	-0.30	-0.48	-0.49	-0.59	-0.53	-0.55	-0.48
Constipation (S)	-0.21	-0.20	-0.49	-0.43	-0.53	-0.45	-0.72	-0.62
Decreased appetite (S)	-0.25	-0.22	-0.41	-0.33	-0.41	-0.31	-0.40	-0.31
Decreased appetite (I)	-0.21	-0.20	-0.41	-0.36	-0.45	-0.35	-0.53	-0.39
Dry mouth (S)	-0.14	-0.14	-0.36	-0.30	-0.51	-0.40	-0.47	-0.35
Dyspnea (S)	-0.11	-0.14	-0.24	-0.28	-0.40	-0.45	-0.55	-0.50
Dyspnea (I)	-0.18	-0.22	-0.22	-0.25	-0.37	-0.35	-0.44	-0.40
Fatigue (S)	-0.30	-0.32	-0.56	-0.61	-0.51	-0.54	-0.53	-0.53
Fatigue (I)	-0.33	-0.30	-0.66	-0.60	-0.72	-0.63	-0.73	-0.65
Insomnia (S)	-0.30	-0.31	-0.45	-0.43	-0.62	-0.54	-0.70	-0.56
Loose stools (F)	-0.02	-0.03	-0.31	-0.31	-0.42	-0.42	-0.58	-0.51
Mouth sores (S)	-0.15	-0.14	-0.39	-0.35	-0.48	-0.39	-0.45	-0.35
Mouth sores (I)	-0.11	-0.12	-0.30	-0.28	-0.44	-0.37	-0.44	-0.36
Nausea (F)	-0.09	-0.10	-0.23	-0.22	-0.28	-0.27	-0.39	-0.34
Nausea (S)	-0.10	-0.14	-0.11	-0.13	-0.25	-0.27	-0.31	-0.30
Neuropathy (S)	-0.28	-0.22	-0.35	-0.27	-0.44	-0.33	-0.57	-0.42
Neuropathy (I)	-0.27	-0.25	-0.32	-0.27	-0.46	-0.37	-0.49	-0.40
Pain (F)	-0.11	-0.10	-0.34	-0.28	-0.51	-0.39	-0.60	-0.45
Pain (S)	-0.20	-0.21	-0.38	-0.38	-0.49	-0.46	-0.44	-0.41
Pain (I)	-0.28	-0.32	-0.47	-0.51	-0.49	-0.52	-0.42	-0.40
Sadness (F)	-0.14	-0.14	-0.36	-0.30	-0.51	-0.40	-0.47	-0.35
Sadness (S)	-0.11	-0.14	-0.24	-0.28	-0.40	-0.45	-0.55	-0.50
Sadness (I)	-0.29	-0.34	-0.42	-0.42	-0.40	-0.40	-0.45	-0.39
Vomiting (F)	0.00	0.00	-0.09	-0.15	-0.25	-0.28	-0.14	-0.14
Vomiting (S)	0.04	0.08	-0.01	-0.01	-0.07	-0.08	-0.11	-0.10

	Maximum of 7 days vs. 1-week recall (N=126*)	Maximum of 14 days vs. 2-week recall (N=118)	Maximum of 21 days vs. 3-week recall (N=109)	Maximum of 28 days vs. 4-week recall (N=93)
Median difference or effect size	Mean Difference	Mean Difference	Mean Difference	Mean Difference
	Effect size	Effect size	Effect size	Effect size
	-0.20	-0.36	-0.45	-0.47
	-0.20	-0.31	-0.39	-0.40

* One patient completed at least one daily report but not the recall report in Week 1, but was included in all subsequent weeks (explaining why 127 patients are included in Table 1).

(F): frequency, (S): severity, (I): interference

Positive (negative) mean differences indicate that the (1-, 2-, 3-, or 4-week) recalled scores were greater (less) than the maximum of the daily scores.

Table 3

Mean differences between average of daily ratings and longer recalled ratings, by recall period, for each PRO-CTCAE item (0–4 scale).

	Average of 7 days vs. 1-week recall (N=126*)		Average of 14 days vs. 2-week recall (N=118)		Average of 21 days vs. 3-week recall (N=109)		Average of 28 days vs. 4-week recall (N=93)	
	Mean Difference	Effect size	Mean Difference	Effect size	Mean Difference	Effect size	Mean Difference	Effect size
Anxiety (F)	0.16	0.17	0.20	0.24	0.33	0.33	0.29	0.28
Anxiety (S)	0.08	0.10	0.13	0.19	0.22	0.28	0.27	0.30
Anxiety (I)	0.00	0.01	0.04	0.06	0.07	0.09	0.13	0.15
Constipation (S)	0.21	0.24	0.27	0.28	0.41	0.38	0.38	0.36
Decreased appetite (S)	0.14	0.14	0.27	0.24	0.41	0.33	0.56	0.42
Decreased appetite (I)	0.09	0.10	0.22	0.23	0.34	0.31	0.49	0.38
Dry mouth (S)	0.12	0.12	0.24	0.23	0.17	0.14	0.32	0.27
Dyspnea (S)	0.08	0.13	0.14	0.19	0.09	0.11	0.15	0.19
Dyspnea (I)	0.01	0.02	0.10	0.14	0.09	0.11	0.09	0.13
Fatigue (S)	0.20	0.24	0.34	0.40	0.49	0.55	0.64	0.65
Fatigue (I)	0.21	0.20	0.29	0.30	0.38	0.37	0.62	0.59
Insomnia (S)	0.16	0.20	0.29	0.35	0.36	0.39	0.38	0.37
Loose stools (F)	0.27	0.34	0.29	0.34	0.33	0.45	0.42	0.43
Mouth sores (S)	0.21	0.23	0.29	0.32	0.45	0.46	0.60	0.56
Mouth sores (I)	0.19	0.23	0.31	0.36	0.41	0.44	0.56	0.51
Nausea (F)	0.06	0.08	0.13	0.16	0.17	0.20	0.14	0.15
Nausea (S)	0.02	0.03	0.12	0.16	0.06	0.08	0.06	0.08
Neuropathy (S)	0.16	0.14	0.49	0.42	0.55	0.43	0.57	0.47
Neuropathy (I)	0.14	0.14	0.45	0.43	0.49	0.44	0.54	0.49
Pain (F)	0.19	0.18	0.35	0.33	0.41	0.38	0.47	0.40
Pain (S)	0.12	0.14	0.20	0.24	0.25	0.28	0.30	0.33
Pain (I)	0.05	0.07	0.10	0.15	0.21	0.27	0.32	0.34
Sadness (F)	-0.08	-0.12	-0.03	-0.06	0.09	0.14	0.17	0.21
Sadness (S)	0.08	0.13	0.14	0.19	0.09	0.11	0.15	0.19
Sadness (I)	0.01	0.02	0.10	0.14	0.09	0.11	0.09	0.13
Vomiting (F)	0.09	0.19	0.15	0.25	0.23	0.36	0.43	0.47
Vomiting (S)	0.11	0.21	0.19	0.29	0.34	0.40	0.49	0.40

	Average of 7 days vs. 1-week recall (N=126)*		Average of 14 days vs. 2-week recall (N=118)		Average of 21 days vs. 3-week recall (N=109)		Average of 28 days vs. 4-week recall (N=93)	
Median difference or effect size	Mean Difference	Effect size	Mean Difference	Effect size	Mean Difference	Effect size	Mean Difference	Effect size
	0.12	0.14	0.24	0.24	0.33	0.33	0.41	0.36

* One patient completed at least one daily report but not the recall report in Week 1, but was included in all subsequent weeks (explaining why 127 patients are included in Table 1).

(F): frequency, (S): severity, (I): interference

Positive (negative) mean differences indicate that the (1-, 2-, 3-, or 4-week) recalled scores were greater (less) than the average of the daily scores.

Table 4

Intraclass correlation coefficients between maximum of daily ratings and longer recalled ratings, by recall period, for each PRO-CTCAE item

PRO-CTCAE Item	Maximum of 7 days vs. 1-week recall (N=126*)	Maximum of 14 days vs. 2-week recall (N=118)	Maximum of 21 days vs. 3-week recall (N=109)	Maximum of 28 days vs. 4-week recall (N=93)
Anxiety (F)	0.71	0.63	0.74	0.74
Anxiety (S)	0.64	0.60	0.70	0.76
Anxiety (I)	0.64	0.65	0.66	0.67
Constipation (S)	0.67	0.67	0.70	0.72
Decreased appetite (S)	0.81	0.82	0.79	0.81
Decreased appetite (I)	0.73	0.67	0.68	0.81
Dry mouth (S)	0.78	0.80	0.79	0.86
Dyspnea (S)	0.75	0.82	0.71	0.72
Dyspnea (I)	0.65	0.75	0.72	0.69
Fatigue (S)	0.68	0.66	0.70	0.64
Fatigue (I)	0.62	0.58	0.61	0.75
Insomnia (S)	0.64	0.75	0.67	0.69
Loose stools (F)	0.66	0.71	0.62	0.68
Mouth sores (S)	0.70	0.76	0.80	0.79
Mouth sores (I)	0.70	0.74	0.74	0.80
Nausea (F)	0.70	0.69	0.73	0.71
Nausea (S)	0.71	0.72	0.77	0.69
Neuropathy (S)	0.82	0.75	0.62	0.73
Neuropathy (I)	0.71	0.75	0.65	0.71
Pain (F)	0.81	0.82	0.84	0.77
Pain (S)	0.78	0.83	0.82	0.77
Pain (I)	0.81	0.78	0.76	0.68
Sadness (F)	0.71	0.75	0.78	0.71
Sadness (S)	0.69	0.67	0.73	0.77
Sadness (I)	0.61	0.62	0.66	0.65
Vomiting (F)	0.54	0.75	0.63	0.69
Vomiting (S)	0.61	0.69	0.82	0.78
Median intraclass correlation	0.70	0.74	0.72	0.72

* One patient completed at least one daily report but not the recall report in Week 1, but was included in all subsequent weeks (explaining why 127 patients are included in Table 1).

(F): frequency, (S): severity, (I): interference

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript