



Published in final edited form as:

Clin Ther. 2016 April ; 38(4): 821–830. doi:10.1016/j.clinthera.2016.03.011.

Methods for implementing patient-reported outcome (PRO) measures of symptomatic adverse events in cancer clinical trials

Ethan Basch, MD, MSc^a, Lauren J. Rogak, MA^b, and Amylou C. Dueck, PhD^c

Ethan Basch: ebasch@med.unc.edu; Lauren J. Rogak: rogakl@mskcc.org; Amylou C. Dueck: dueck@mayo.edu

^aCancer Outcomes Research Program, Lineberger Comprehensive Cancer Center, University of North Carolina, 170 Manning Drive, Chapel Hill, NC 27599, USA

^bHealth Outcomes Research Group, Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, 485 Lexington Avenue, New York, NY 10017, USA

^cSection of Biostatistics, Mayo Clinic, 13400 E. Shea Blvd, Scottsdale, AZ 85259, USA

Abstract

Purpose—There is increasing interest to use patient-reported outcome (PRO) measures to evaluate symptomatic adverse events (AEs) in cancer treatment trials. However, there are currently no standard recommended approaches for integrating patient-reported AE measures into trials.

Methods—Approaches are identified from prior trials for selecting AEs for solicited patient-reporting; administering patient-reported AE measures; and analyzing and reporting results.

Findings—Approaches for integrating patient-reported AE measures into cancer trials generally combine current standard methods for clinician-reported AEs as well as established best practices for employing PRO measures. Specific AEs can be selected for a PRO questionnaire based on common and expected reactions in a given trial context, derived from literature review and qualitative/mixed methods evaluations and should be the same set administered across all arms of a trial. A mechanism for collecting unsolicited patient-reported AEs will also ideally be included. Patients will preferably report at baseline and end of active treatment as well as on a frequent standardized schedule during active treatment, such as weekly from home, with a recall period corresponding to the frequency of reporting (e.g., past 7 days). Less frequent reporting may be considered after an initial intensive monitoring period for trials of prolonged treatments and during long-term follow up. Electronic PRO data collection is preferred. Backup data collection for missed PRO reports is advisable to boost response rates. Analysis can employ a combination of approaches to AE and PRO data. If a high proportion of patients is experiencing baseline symptoms, systematic subtraction of these from on-study AEs should be considered to improve

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

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

Conflict of Interest statement

The authors have no commercial interests in any questionnaire system and report no financial conflicts of interest.

reporting of symptoms related to treatment. More granular longitudinal analyses of individual symptoms can also be useful.

Implications—Methods are evolving for integrating patient-reported symptomatic AEs into cancer trials. These methods are expected to further evolve as more data from trials become available.

Introduction

Patient-reported outcomes (PROs) are considered the gold standard for assessing symptoms in clinical research¹. Although historically, PRO measures have been used largely to measure symptoms related to a disease process (e.g., pain related to bone metastases) or as outcomes in symptom management trials, there is increasing interest to use PRO measures to assess symptomatic adverse events (AEs) in cancer treatment trials (e.g., nausea related to platinum agents).² Cancer therapies are often associated with significant symptomatic AEs, and rigorous collection of information about these AEs is an essential component of understanding treatment characteristics and impact on patients.³

The current standard approach to AE reporting in cancer treatment trials is based on clinician reporting using the Common Terminology Criteria for Adverse Events (CTCAE), which is maintained by the U.S. National Cancer Institute (NCI)⁴. For each AE within CTCAE, clinical grades can range from 0 to 5 representing severity levels of none, mild, moderate, severe, life-threatening, or death, though the maximum possible grade and specific grade level description/criteria can vary depending on the AE. Multiple studies have demonstrated that this approach misses up to half of AEs compared to patient self-report, and that PRO measures improve the detection and precision of AE measurement^{5,6,7}. The U.S. Food and Drug Administration (FDA) has suggested the use of PRO measures for adverse event measurement in oncology drug development².

There are a number of potential uses of patient-reported AE measures in cancer clinical trials. Uses include: 1) In early-phase trials to generate initial information about AEs, and to support selection of tolerable dose levels and schedules; 2) in pivotal trials to measure baseline symptoms and to characterize symptomatic adverse event profiles; 3) in comparative trials to provide data for comparing tolerability between treatments; and 4) in post-marketing studies for long-term and broad population safety surveillance.

There is currently no standard recommended approach for integrating patient-reported symptomatic AEs into cancer treatment trials. The purpose of this paper is to describe approaches that have been used in prior trials. In general, the approaches described in this paper combine standard methodologies for employing the CTCAE as well as established best practices for employing PRO measures in cancer clinical trials. This area is rapidly evolving and as more data become available, increasingly evidence-based recommendations are expected to become available.

Selection of adverse events (outcomes) for measurement

Ideally, a set of symptomatic adverse events should be identified and specified *a priori* in a trial protocol for systematic assessment. The same set of symptomatic adverse events should be assessed in all arms of the trial to allow for between-arm comparisons. This set may be derived from three potential sources.

First, a “core” set of common symptomatic adverse events may be included regardless of the planned treatment(s). In oncology such a core set has been previously identified through an NCI-supported consensus process, and includes anorexia, anxiety, cognitive disturbance, constipation, depression, diarrhea, dyspnea, fatigue, insomnia, nausea, neuropathy, and pain⁸. This list is not specific to cancer type or therapy type and may include some disease-related rather than treatment-related symptoms. Therefore, a subset or variation may be considered, depending on the trial and available evidence about the population and treatment characteristics. Notably, if a broader health-related quality of life questionnaire is also being administered in a given trial that includes some of the “core” symptoms such as anxiety or depression, these may be collected with that tool and not with the patient-reported AE questionnaire to avoid redundancy. A goal of thoughtful AE selection for a PRO questionnaire is to determine a parsimonious set of the most salient symptomatic adverse events, maximizing relevance and minimizing burden and duplication. Beyond using a structured PRO questionnaire for collecting solicited AEs, a mechanism for allowing free text unsolicited AE reporting by patients should be considered. Second, established data about expected adverse effects of particular treatments or classes of treatments can be considered (e.g., myalgias with aromatase inhibitors or neuropathy with platinum agents). Third, qualitative or mixed methods work during early-phase research to identify potential symptomatic AEs for subsequent assessment can be highly informative, for example, in a nonrandomized expansion cohort.

Selection of measures

In general, for adverse event assessment, a limited number of questions is necessary for each symptom of interest, as AE reporting is intended to screen rather than provide in-depth assessment. A variety of tools have been developed to assess patient-reported symptoms in cancer trials that offer a wide variety of symptoms that may occur as treatment-related adverse events⁹.

Measures which include items that mirror selected CTCAE items^{10,11,12,13} have been used to compare patient and physician grading or assess the usefulness of PRO measures in clinical settings. A number of measures that have been used to measure adverse events in cancer clinical trials were developed to assess disease-related symptoms rather than adverse events explicitly, or to evaluate broader domains of quality of life⁹. Recently, the NCI developed the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) with the explicit purpose to evaluate symptomatic adverse events^{14,15,16}. The PRO-CTCAE is a library of 124 questions that measure 78 specific symptomatic adverse events that are common in oncology. The PRO-CTCAE was designed and evaluated to enable investigators to pick and choose questions corresponding to

whichever specific AEs are pertinent to a given trial, for example based on the approach listed above. For each AE within PRO-CTCAE, items measure up to three attributes depending on the AE, with possible attributes of frequency (with response options of “never”, “rarely”, “occasionally”, “frequently”, or “almost constantly”), severity (“none”, “mild”, “moderate”, “severe”, or “very severe”), interference with usual or daily activities (“not at all”, “a little bit”, “somewhat”, “quite a bit”, or “very much”), amount (“not at all”, “a little bit”, “somewhat”, “quite a bit”, or “very much”), or presence (“no” or “yes”)¹⁴.

Study design and data collection

Frequency of patient-reporting

Patient-reported AEs can be collected systematically via a pre-populated questionnaire at baseline, regularly during active treatment, and at end of treatment, with consideration during post-treatment (long-term) follow-up. Ideally, patients will self-report weekly during active treatment to ensure comprehensive capture of AE data. If weekly reporting is not feasible, then reporting at least once per cycle is desirable, with collection of data at a time point when AEs are not expected to be resolved. In a trial of prolonged treatment where AEs are expected to stabilize after an initial period, an approach can be considered with more frequent reporting initially with greater spacing later on (e.g., weekly reporting for 6 months, followed by monthly or quarterly reporting subsequently during active treatment). Similarly, for post-treatment follow-up, if continued patient-based AE assessments are of interest to characterize late toxicities, evaluations may be spaced out (e.g., every 3–6 months for up to 3 years). Expected timing of worst side effects relative to dosing may also inform the frequency of scheduled patient-reporting (i.e., the schedule of assessments may differ for agents with continuous oral dosing versus intravenous chemotherapy administered on a limited number of days each cycle). A baseline patient-reported AE assessment is essential to enable understanding of those symptoms that are treatment emergent versus pre-existing, as discussed below.

Recall period

Another consideration related to timing of administration is the recall period of the PRO measure. Many measures include a 1-week recall, although the PRO-CTCAE, for example, includes flexibility to extend this to 2-week, 3-week, or 4-week recall. Evidence suggests that patients are able to adequately recall a variety of symptoms in the prior week with results similar to daily reporting^{17,18}. With longer recall periods there is some information loss likely related to memory degradation, although correlations with daily reports remain relatively high with up to 4-week recall¹⁷. Given the potential differences in the amount of memory degradation across recall periods, recall period should be standardized across arms in multi-arm comparative trials.

When selecting a recall period in a given trial, the balance between comprehensive coverage of elapsed time must be balanced with study logistics. For example, if remote home reporting is feasible, then weekly reporting with a 1-week recall is preferable (particularly if symptoms are expected to fluctuate or change from week to week). This is the authors’ preferred approach. However, if patients are completing questionnaires at clinic visits every

3-week cycle, then a 3-week recall may be desirable with an understanding that there may be some information loss. There may also be certain circumstances in which capture of daily symptomatic AE reports may be appropriate. Like weekly reporting from home, daily at-home reporting is potentially feasible with the use of field-based electronic data collection. One of the clear advantages of remote electronic data capture is that the data can be uploaded to a server or dataset in real time and alerts can be sent to site investigators if a particularly significant and/or severe AE is reported (see next section on modes of data collection, as well as the section on sharing patient-reported AEs with clinicians).

Mode of data collection

Different modes of data collection are available for PRO measures, including paper and electronic formats (e.g., Web-based, handheld device application-based, and automated telephone system-based). Electronic modes are preferable whenever feasible as they enable systematically timed reporting from home between visits as well as automated reminders to patients, automated alerts to investigators (see section on sharing patient-reported AEs with clinicians), and real-time monitoring of compliance. Prior research demonstrates that results are comparable across various modes of administration^{19,20,21}, supporting the use of previously tested modes within a given trial based on convenience and patient preference (see Eremenco et al.²² for design and statistical analysis considerations when mixing modes). An option of a telephone-based approach is desirable, particularly for patients with limited literacy or tactile function, whereas paper or Web-based is preferable with hearing impairment. Assuring availability of a questionnaire in all of the anticipated languages spoken by trial participants should also be considered up front.

Backup data collection

Patient compliance with AE self-reporting can be enhanced substantially with regular reminders to self-report, and backup data collection for those who do not self-report on schedule. This is easier to accomplish with electronic data collection, as automated electronic reminders and notifications for non-compliance can be set up. For example, patients who are self-reporting weekly via Web or automated telephone system can receive an auto-generated reminder by email/call to self-report on a given day. For non-responders, repeat reminders can be automatically triggered up to twice daily, and after 3 days a human backup data manager can call the patient to administer the questions directly.

Paper-based reporting is a less desirable but frequently used option and can have high compliance rates. If paper reporting is conducted at clinic visits, compliance is dependent on local site research staff remembering to provide the questionnaire to patients. If patients miss a visit, the report will be missed, leading to potential informative missingness. Staff can call patients to administer the questionnaire verbatim by telephone in such cases. Between-visit reporting via paper is also an option by providing a PRO questionnaire(s) to patients; however, this approach is generally not recommended as the actual date of completion for each questionnaire cannot be tracked and patients may complete questionnaire(s) retrospectively just prior to turning in the questionnaire(s). If at-home paper-based reporting is unavoidable, regular reminders at visits and calls from staff can be initiated to remind patients to complete assessments on the intended dates. Provision of self-addressed stamped

envelopes for patients to mail completed questionnaires back to clinic may also be considered. When patients return to clinic, if reports have not been completed they can be administered by staff. If patients become too ill to self-report, assistance by a family member, friend, or caregiver serving as an objective reporter or recorder of the patient's responses may be employed but should be demarcated in the dataset to allow for proper handling in statistical analysis (e.g., as a covariate in statistical models and/or sensitivity analyses excluding such data to assess the impact).

Sharing patient-reported adverse events with clinicians

There is an option to share patient-reported AEs with physicians and nurses at the point of care to inform clinical management as well as clinician adverse event grading. With electronic reporting, automated notifications for severe or worsening symptomatic adverse events can be triggered, and full reports can be printed or viewed at visits. This approach has been found to be feasible and perceived as useful by clinicians, and most notifications are seen as clinically actionable²³. Moreover, the sharing of patient-reported adverse event scores at the point of care can improve the precision of clinician-reported adverse events⁶. This approach simply formalizes and systematizes the existing method for symptomatic adverse event detection in clinical trials, which is primarily based on interviews of patients at visits. Clinicians are responsible for understanding and documenting their patients' adverse events in trials, and this approach improves that function. Finally, informing patients that their self-reported AEs will be shared with their clinicians to improve care may potentially improve patient compliance with reporting, as patients may view the patient-reported AE assessment not just as research data collection but as a way to engage the clinician in their own treatment and health status. However, the potential benefits of systematically sharing patient-reported AEs with clinicians must be balanced with clinical trial logistics and potential impacts of sharing these data on conduct of the ongoing study (e.g., inadvertent unblinding in a blinded trial or change in enthusiasm for continued site enrollment) on a case-by-case basis.

For example, in a phase 2 lung cancer trial, patients self-reported adverse events at clinic visits via tablet computers, and this information was shared with physicians and nurses in real time in a clinician interface which allowed the clinician to agree or disagree/reassign the patient-reported grades²³. Clinicians agreed with the patient-reported AEs in 93% of cases, although agreement levels were lower for high-grade AEs.

Future directions

The field of patient-based AE reporting in cancer trials remains evolving in multiple aspects of design, statistical analysis, and reporting strategies. Based on a limited number of clinical trials which have incorporated patient-reported AE measurement to date, statistical analyses have generally combined approaches developed for the analysis of traditional clinician-based AE data with strategies developed for PRO measures. As additional studies are completed, optimal statistical analysis and reporting strategies will likely continue to combine the approaches used for traditional clinician-based AE data and PRO data, taking into consideration two key differences between PRO data and traditional clinician-based AE

data: first, that patients better detect baseline symptoms than clinicians²⁸, which can be considered in analyses; and second that patient-reporting can generate more frequent between-visit reports allowing for granular longitudinal analyses.

The standard approach for CTCAE data analysis often involves tabulation of the maximum grade post-baseline per patient per adverse event. In multi-arm trials, this is conducted by arm. For each AE, the incidence of grade 1 or higher is computed (i.e., grade >0), and separately the incidence of grade 3 or higher is computed. In multi-arm trials, rates are often compared between arms using Fisher's exact tests, chi-squared tests, or similar method. Additional analysis may involve similar tabulation of only adverse events deemed at least possibly related to study treatment by the treating clinician, or tabulation of treatment-emergent adverse events. These data are commonly displayed descriptively (with or without p-values) in tabular format in publications. Patient-reported AEs may be analyzed and displayed similarly, either in a stand-alone patient-reported AE table, or in a combined table which shows both patient-reported and clinician reported AEs (example shown in Table 1).

Alternative statistical analysis strategies have been developed but not commonly employed for CTCAE-based data such as summary statistics other than the maximum grade post-baseline and multivariate and time-to-event analyses^{24,25,26}. Use of alternative approaches may become more common or new approaches may be developed in the future with the increased testing of continuous-dosing oral targeted therapies taken over longer periods of time with potentially lower grade toxicities relative to traditional chemotherapies²⁷.

In addition, histograms may be used to provide more granular depictions of the full spectrum of scores. These can be used to summarize the distribution of multiple post-baseline AEs during treatment (example shown in Figure 1), or to show the distribution of grades or scores of individual AEs at each assessment time point in a longitudinal fashion (example shown in Figure 2).

When there are substantial rates of baseline symptoms, analyses may be adjusted to remove these from results to avoid misattribution of pre-existing symptoms to study treatments. Patients with advanced cancers frequently enter trials with baseline symptoms that are not reported by clinicians (for example, see Figure 2, Panel 2b), but are detected by patient-reporting, which has greater precision²⁸. Indeed, many cancer clinical trials do not include baseline CTCAE reporting. When clinicians are aware of baseline symptoms, they generally consider these during their subsequent CTCAE grading implicitly by not reporting these as adverse events unless the symptom worsens (e.g., pre-existing dyspnea in a lung cancer trial). In PRO data analyses generally, an approach frequently used to adjust for baseline scores is to tabulate change from baseline. However, for adverse event reporting a different approach is more synonymous with clinician CTCAE grading and may improve the attribution of symptoms to specific treatments. Specifically, for any given patient, the worst adverse event during treatment is tabulated only for adverse events that are worse than the baseline score. For example, if a patient had nausea at baseline with a magnitude score of 2, and his or her worst post-baseline score was 1 or 2, then no AE would be tabulated for that patient (i.e., a score of 0). However, if that patient's worst post-baseline score was 3 or

higher, then that post-baseline score would be tabulated for that patient. This ‘baseline subtraction’ method is used in the data shown in Table 1.

It is well established that patient and clinician symptom reports are discrepant, with clinicians generally under-reporting the incidence and magnitude of symptoms compared to patients. As previously noted, sharing patient reports with clinicians at the point of care improves the alignment of patient and clinician grades, but discrepancies persist. Patient-reported and clinician-reported AEs can be presented together with an understanding that they are discrepant and may capture complementary aspects of adverse event phenomena²⁹. This can be done either in a combined AE table as demonstrated in Table 1, or with separate tables for clinician and patient-reported AEs. A second option is to report only the patient-reported AEs where symptomatic AEs are concerned, and to report clinician-reported CTCAE only for non-symptomatic AEs. Attempts to reconcile discrepant patient and clinician AE reports after the fact are likely to be logistically infeasible given that the patient report is its own source documentation. While further information may be attainable from the medical record with regard to the clinician’s reported grade, any further information about the patient’s reported score would require contact with the patient who may have terminated follow-up or died, with recall of the event being subject to memory degradation if the patient was in fact reachable.

Other statistical analysis considerations for patient-reported adverse event data include interpretation in single-arm trials, handling missing data, and multiplicity associated with statistical hypothesis testing (e.g., between-arm comparisons in multi-arm trials). Statistical analyses for clinician-reported CTCAE data in part address missing data through the use of summary measures (e.g., tabulating the worst grade post-baseline). While a similar approach can be employed for patient-reported AEs, a major caveat is that patients who are ill or hospitalized may miss self-reports, therefore yielding informative missingness. Various statistical methods including imputation or model-based approaches have been applied to PRO-based efficacy endpoints in clinical trials and are available for patient-reported AE analysis, either as part of the primary statistical approach or for use in sensitivity analyses. One intriguing yet still untested approach to missing patient-reported AE data is the use of the clinician-reported AE data either as values for direct imputation or as auxiliary data for model-based multiple imputation, weighting or stratification in statistical modeling, joint auxiliary and outcome modeling, or investigating missing data mechanisms³⁰. As noted above, use of electronic PRO data collection and efforts to recover missing self-reports via a variety of backup approaches are recommended to minimize missing data, and can raise compliance rates into the 90% range.³¹

Relative to multiplicity, CTCAE data tabulated by arm in multi-arm trials are often presented descriptively (i.e., without p-values for between-arm comparisons) as the goal of AE reporting is detection of potential safety signals and description of side effect profiles for a given treatment, not formal statistical hypothesis testing. However, as noted above, p-values may be used, and the precision of patient-reported AE assessment may yield a higher number of statistically significant differences between study arms than traditional CTCAE reporting. The FDA has noted that patient-reported AE data may be useful when presented descriptively², although optimal methods for considering multiplicity adjustment when

formal statistical hypothesis testing beyond descriptive reporting is carried out has yet to be determined. While various methods for multiplicity adjustment and the statistical issues raised above exist for traditional clinician-based AE data and general PRO-based endpoints, these statistical analysis and reporting techniques will likely evolve as data from additional trials systematically incorporating patient-reported AEs become available.

Conclusion

Asking patients to systematically self-report symptomatic AEs in cancer treatment trials has the potential to improve the quality of AE detection, while enabling greater patient engagement in clinical research. Assessment tools and methodologies already exist that can be employed in trials, including NCI's PRO-CTCAE. This paper outlines some of the key issues to consider when integrating patient self-report into the overall assessment of AEs in clinical research and shows that, by complementing the traditional clinician-based AE reporting system (i.e., CTCAE) with patient-reported AE data, a much clearer understanding of tolerability is gained and the evaluation of investigational anti-cancer agents is enriched. We anticipate optimal integration, statistical analysis, and reporting strategies will evolve over time as patient-based adverse event reporting becomes more common.

Acknowledgments

The views expressed in this commentary reflect the opinions of the authors based on published literature and their own experiences using clinician-based and patient-based adverse event reporting in cancer clinical trials and practice. This commentary does not reflect formal recommendations on use of any particular questionnaire system for assessing adverse events, and does not reflect recommendations by the U.S. Food and Drug Administration or National Cancer Institute.

The authors conducted studies using the STAR system which was developed at Memorial Sloan Kettering Cancer Center, and served as investigators under contract or subcontracts from the U.S. National Cancer Institute to Memorial Sloan Kettering Cancer Center to develop the PRO-CTCAE system.

References

1. U.S. Department of Health and Human Services, Food and Drug Administration. [last accessed 01/21/16] Guidance for industry: Patient-reported outcomes measures: Use in medical product development to support labeling claims. 2009 Dec. Available at <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf>
2. Kluetz PG, Slagle A, Papadopoulos E, Johnson LL, Donoghue M, Kwitkowski VE, Chen WH, Sridhara R, Farrell AT, Keegan P, Kim G, Pazdur R. Focusing on Core Patient-Reported Outcomes in Cancer Clinical Trials: Symptomatic Adverse Events, Physical Function, and Disease-Related Symptoms. *Clin Cancer Res*. 2016 Jan 12. [Epub ahead of print] PubMed PMID: 26758559.
3. Basch E. The missing voice of patients in drug-safety reporting. *N Engl J Med*. 2010 Mar 11; 362(10):865–869. [PubMed: 20220181]
4. 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
5. Di Maio M, Gallo C, Leighl 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 Mar 10; 33(8):910–915. PubMed PMID: 25624439. [PubMed: 25624439]

6. 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 event reporting by patients vs clinicians: relationships with clinical outcomes. *J Natl Cancer Inst.* 2009 Dec 2; 101(23):1624–1632. PubMed PMID: 19920223; PubMed Central PMCID: PMC2786917. [PubMed: 19920223]
7. 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 Nov; 7(11):903–909. PubMed PMID: 17081915. [PubMed: 17081915]
8. 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) pii: dju129. Print 2014 Jul. Review. PubMed PMID: 25006191; PubMed Central PMCID: PMC4110472.
9. Basch E, Abernethy AP, Mullins CD, Reeve BB, Smith ML, Coons SJ, Sloan J, Wenzel K, Chauhan C, Eppard W, Frank ES, Lipscomb J, Raymond SA, Spencer M, Tunis S. Recommendations for incorporating patient-reported outcomes into clinical comparative effectiveness research in adult oncology. *J Clin Oncol.* 2012 Dec 1; 30(34):4249–4255. [PubMed: 23071244]
10. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, Rogak L, Bennett AV, Dueck AC, Atkinson TM, Chou JF, Dulko D, Sit L, Barz A, Novotny P, Fruscione M, Sloan JA, Schrag D. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. *J Clin Oncol.* 2015 Dec 7. pii: JCO630830. [Epub ahead of print] PubMed PMID: 26644527.
11. Cirillo M, Venturini M, Ciccarelli L, Coati F, Bortolami O, Verlato G. Clinician versus nurse symptom reporting using the National Cancer Institute-Common Terminology Criteria for Adverse Events during chemotherapy: results of a comparison based on patient's self-reported questionnaire. *Ann Oncol.* 2009 Dec; 20(12):1929–1935. Epub 2009 Jul 17. PubMed PMID: 19622510. [PubMed: 19622510]
12. Novello S, Capelletto E, Cortinovis D, Tiseo M, Galetta D, Valmadre G, Casartelli C, Rapetti SG, Rossi A. Italian multicenter survey to evaluate the opinion of patients and their reference clinicians on the "tolerance" to targeted therapies already available for non-small cell lung cancer treatment in daily clinical practice. *Transl Lung Cancer Res.* 2014 Jun; 3(3):173–180. PubMed PMID: 25806297; PubMed Central PMCID: PMC4367690. [PubMed: 25806297]
13. Montemurro F, Mittica G, Cagnazzo C, Longo V, Berchiolla P, Solinas G, Culotta P, Martinello R, Foresto M, Gallizioli S, Calori A, Grasso B, Volpone C, Bertola G, Parola G, Tealdi G, Giuliano PL, Aglietta M, Ballari AM. Self-Evaluation of Adjuvant Chemotherapy-Related Adverse Effects by Patients With Breast Cancer. *JAMA Oncol.* 2015 Dec.23:1–8. [Epub ahead of print] PubMed PMID: 26720497.
14. 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) pii: dju244. Print 2014 Sep. PubMed PMID: 25265940; PubMed Central PMCID: PMC4200059.
15. 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):257–269. Epub 2013 Jul 20. PubMed PMID: 23868457; PubMed Central PMCID: PMC3896507. [PubMed: 23868457]
16. 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; 1(8):1051–1059. PubMed PMID: 26270597. [PubMed: 26270597]
17. 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):146–157. [PubMed: 18455312]
 18. Schneider S, Broderick JE, Junghaenel DU, Schwartz JE, Stone AA. Temporal trends in symptom experience predict the accuracy of recall PROs. *J Psychosom Res.* 2013 Aug; 75(2):160–166. Epub 2013 Jul 6. PubMed PMID: 23915773; PubMed Central PMCID: PMC3740272. [PubMed: 23915773]
 19. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health.* 2008 Mar-Apr; 11(2):322–333. [PubMed: 18380645]
 20. Muehlhausen W, Doll H, Quadri N, Fordham B, O'Donohoe P, Dogar N, Wild DJ. Equivalence of electronic and paper administration of patient-reported outcome measures: a systematic review and meta-analysis of studies conducted between 2007 and 2013. *Health Qual Life Outcomes.* 2015 Oct 7. 13:167. PubMed PMID: 26446159; PubMed Central PMCID: PMC4597451. [PubMed: 26446159]
 21. 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 PMID: 26892667; PubMed Central PMCID: PMC4759776. [PubMed: 26892667]
 22. Eremenco S, Coons SJ, Paty J, Coyne K, Bennett AV, McEntegart D. ISPOR PRO Mixed Modes Task Force. PRO data collection in clinical trials using mixed modes: report of the ISPOR PRO mixed modes good research practices task force. *Value Health.* 2014 Jul; 17(5):501–516. PubMed PMID: 25128043. [PubMed: 25128043]
 23. Basch E, Wood WA, Schrag D, Sima CS, Shaw M, Rogak LJ, Kris MG, Shouery M, Bennett A, Atkinson T, Pietanza MC. Feasibility and clinical impact of sharing patient-reported symptom toxicities and performance status with clinical investigators during a phase 2 cancer treatment trial. *Clin Trials.* 2015 Nov 4. pii: 1740774515615540. [Epub ahead of print].
 24. Trotti A, Pajak TF, Gwede CK, Paulus R, Cooper J, Forastiere A, Ridge JA, Watkins-Bruner D, Garden AS, Ang KK, Curran W. TAME: development of a new method for summarising adverse events of cancer treatment by the Radiation Therapy Oncology Group. *Lancet Oncol.* 2007 Jul; 8(7):613–624. PubMed PMID: 17543584. [PubMed: 17543584]
 25. Klingenberg B, Agresti A. Multivariate extensions of McNemar's test. *Biometrics.* 2006 Sep; 62(3):921–928. PubMed PMID: 16984337. [PubMed: 16984337]
 26. van der Pal HJ, van Dalen EC, van Delden E, van Dijk IW, Kok WE, Geskus RB, Sieswerda E, Oldenburger F, Koning CC, van Leeuwen FE, Caron HN, Kremer LC. High risk of symptomatic cardiac events in childhood cancer survivors. *J Clin Oncol.* 2012 May 1; 30(13):1429–1437. Epub 2012 Apr 2. PubMed PMID: 22473161. [PubMed: 22473161]
 27. Thanarajasingam G, Hubbard JM, Sloan JA, Grothey A. The Imperative for a New Approach to Toxicity Analysis in Oncology Clinical Trials. *J Natl Cancer Inst.* 2015 Aug 1. 107(10) pii: djv216. Print 2015 Oct. PubMed PMID: 26232762.
 28. Atkinson TM, Satele DV, Sloan JA, Mehedint D, Lafky JM, Basch EM, Dueck AC. Comparison between clinician and patient-reporting of baseline (bl) and post-bl symptomatic toxicities in cancer cooperative group clinical trials (NCCTG N0591 [Alliance]). *J Clin Oncol.* 2015 May 20. 33(15)
 29. 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 Dec 21; 103(24):1851–1858. [PubMed: 22157640]

30. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. *Stat Methods Med Res.* 2014 Oct; 23(5):440–459. Epub 2013 Feb 19. PubMed PMID: 23427225. [PubMed: 23427225]
31. Movsas B, Hunt D, Watkins-Bruner D, Lee WR, Tharpe H, Goldstein D, Moore J, Dayes IS, Parise S, Sandler H. Can electronic web-based technology improve quality of life data collection? Analysis of Radiation Therapy Oncology Group 0828. *Pract Radiat Oncol.* 2014 May-Jun;4(3): 187–191. Epub 2013 Sep 16. PubMed PMID: 24766686; PubMed Central PMCID: PMC4007315. [PubMed: 24766686]

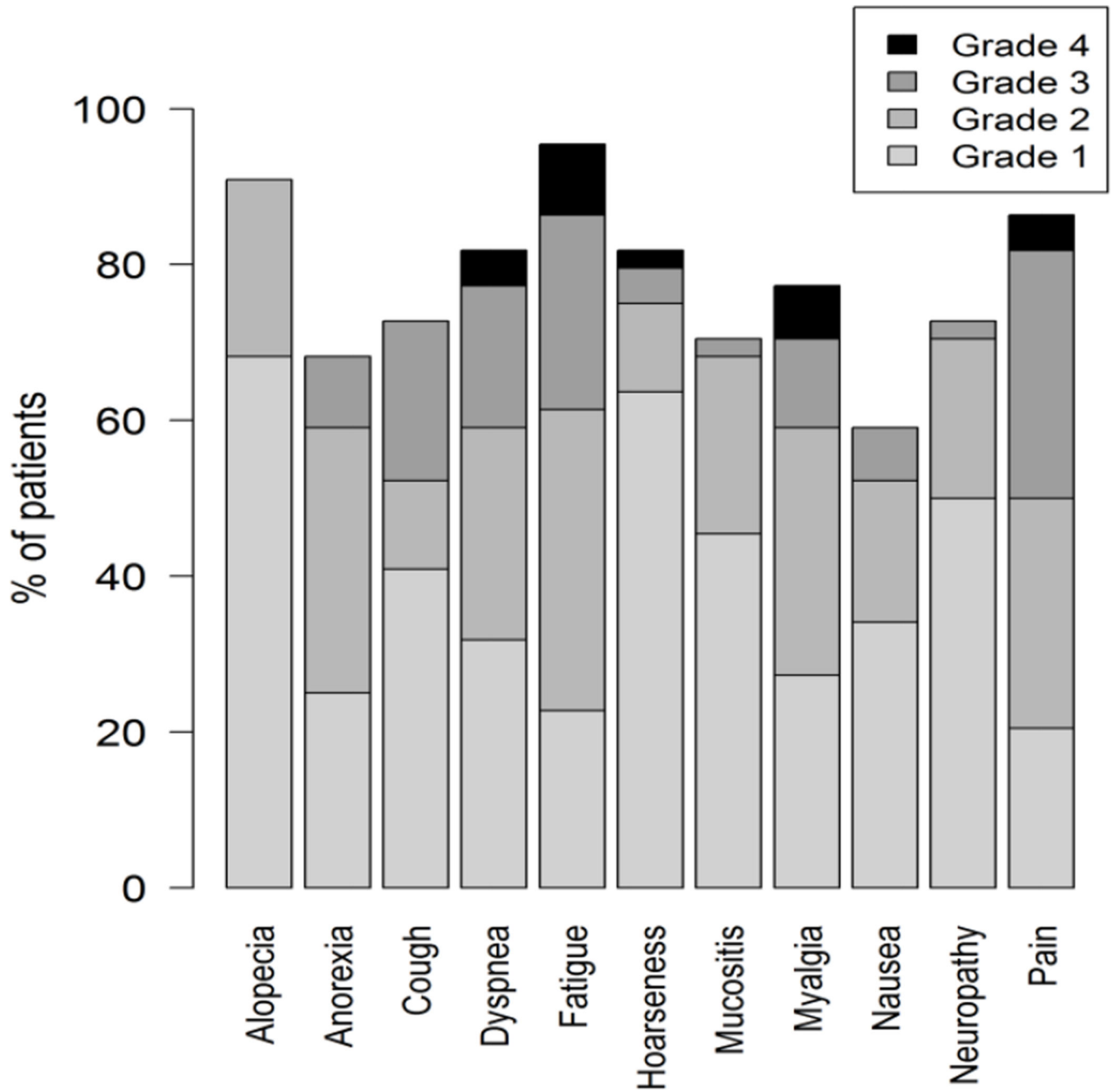
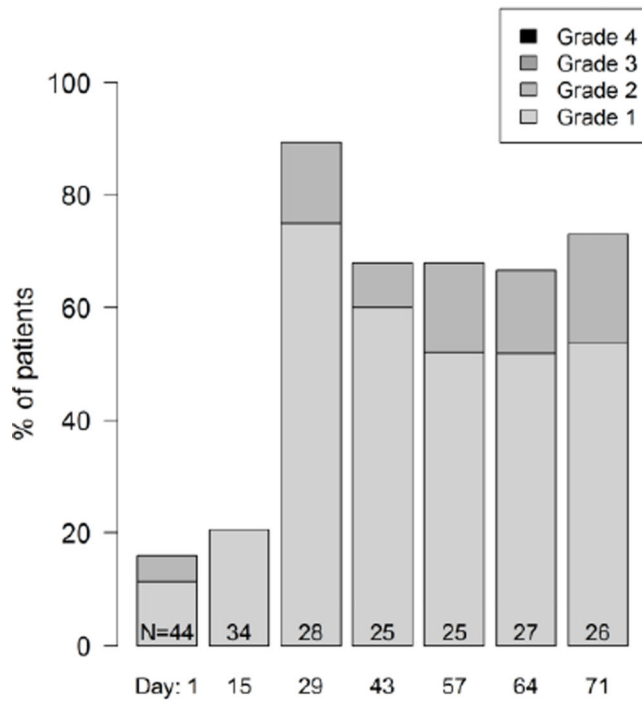
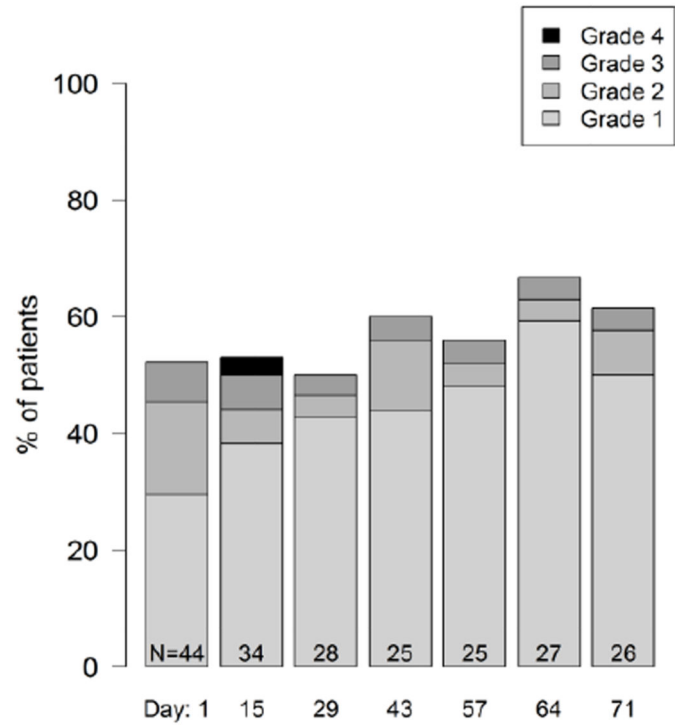


Figure 1. Example graphical display showing the distribution of maximum post-baseline grades reported by patients during treatment in a phase 2 cancer clinical trial (Clin Trials. 2015 Nov 4. pii: 1740774515615540. PMID:26542025). Grades are based on items modified for patient use from the Common Terminology Criteria for Adverse Events (CTCAE).

a. Alopecia



b. Dyspnea



c. Mucositis

d. Nausea

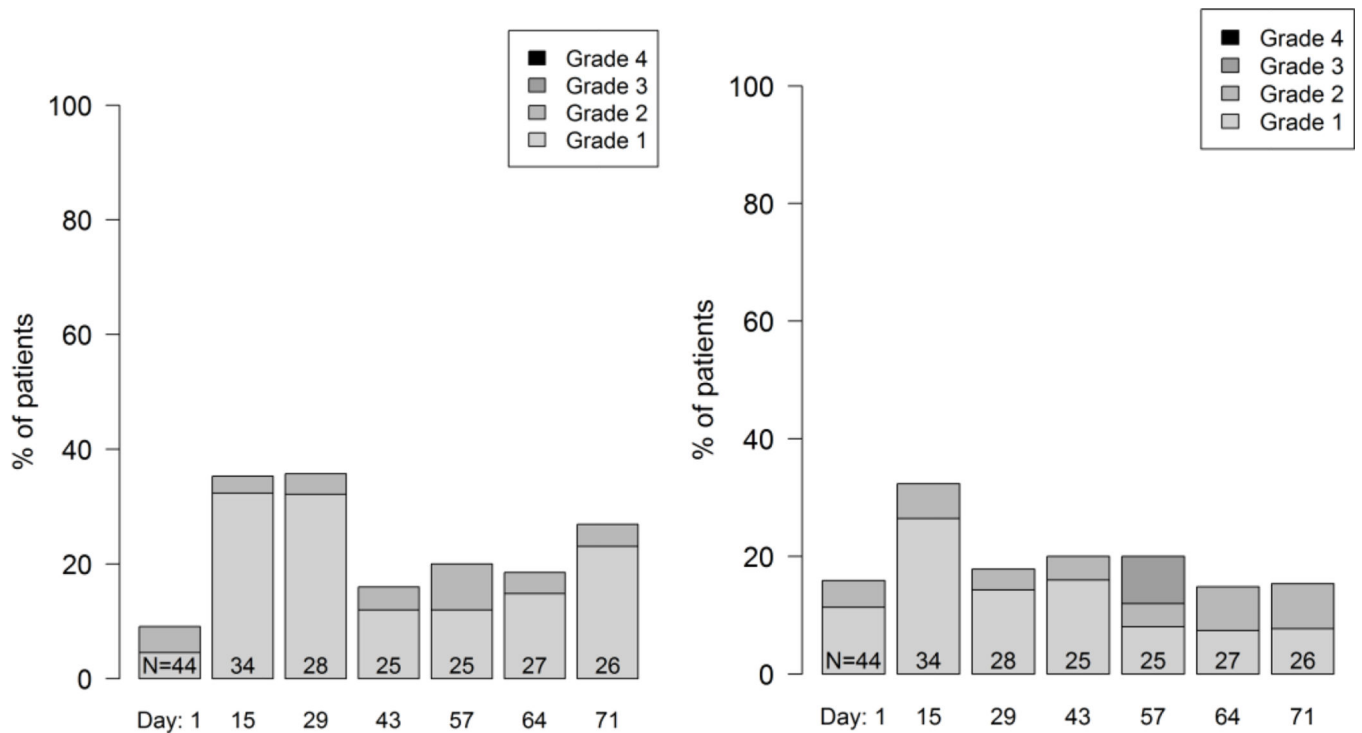


Figure 2. Example longitudinal graphical displays showing the distribution of patient grade levels for four symptomatic adverse events at time points during treatment in a phase 2 cancer clinical trial (Clin Trials. 2015 Nov 4. pii: 1740774515615540. PMID:26542025). Grades are based on items modified for patient use from the Common Terminology Criteria for Adverse Events (CTCAE). Few patients experienced alopecia at baseline (Day 1) and this number markedly increased at Day 28. In contrast, about half of patients had dyspnea at baseline and this proportion did not change substantially, although the distribution of patient-reported grades improved during the trial, suggesting improvement. Mucositis and nausea both increased at Day 15 then improved, suggesting symptomatic control with supportive measures, and suggests the need for earlier preventive measures for mucositis and nausea with the study regimen.

Example table showing the proportion of patients experiencing any grade level, and high-grade level (grade >3) of post-baseline symptomatic adverse events, based on clinician-report and patient-report, using the baseline grade subtraction method for the patient-report. Data are simulated based on observed patterns of clinician-reported and patient-reported grades for patients enrolled in actual clinical studies. Grades are based on Common Terminology Criteria for Adverse Events (CTCAE). This table is provided as illustrative of format only. No information regarding relationship between clinician-reporting and patient-reporting of adverse events should be drawn from this table.

Table 1

Adverse Event	Any Grade		High Grade*		P-value [†]	
	Arm A (N=100)	Arm B (N=100)	Arm A (N=100)	Arm B (N=100)		
Anorexia	<i>Clinician</i>	43 (43%)	36 (36%)	3 (3%)	2 (2%)	0.39
	<i>Patient</i>	57 (57%)	40 (40%)	11 (11%)	8 (8%)	
Constipation	<i>Clinician</i>	27 (27%)	23 (23%)	1 (1%)	--	0.62
	<i>Patient</i>	48 (48%)	41 (41%)	12 (12%)	7 (7%)	
Diarrhea	<i>Clinician</i>	8 (8%)	14 (14%)	1 (1%)	--	0.26
	<i>Patient</i>	39 (39%)	12 (12%)	4 (4%)	--	
Nausea	<i>Clinician</i>	20 (20%)	7 (7%)	--	--	0.01
	<i>Patient</i>	66 (66%)	9 (9%)	9 (9%)	1 (1%)	
Sensory neuropathy	<i>Clinician</i>	54 (54%)	29 (29%)	7 (7%)	4 (4%)	<0.001
	<i>Patient</i>	59 (59%)	51 (51%)	3 (3%)	--	

* Grade >3 based on Common Terminology Criteria for Adverse Events (CTCAE).

[†]Based on Fisher's exact test comparing rate of grade >0 between arms.