

Adverse Symptom Event Reporting by Patients vs Clinicians: Relationships With Clinical Outcomes

Ethan Basch, Xiaoyu Jia, Glenn Heller, Allison Barz, Laura Sit, Michael Fruscione, Mark Appawu, Alexia Iasonos, Thomas Atkinson, Shari Goldfarb, Ann Culkin, Mark G. Kris, Deborah Schrag

- Background** In cancer treatment trials, the standard source of adverse symptom data is clinician reporting by use of items from the National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE). Patient self-reporting has been proposed as an additional data source, but the implications of such a shift are not understood.
- Methods** Patients with lung cancer receiving chemotherapy and their clinicians independently reported six CTCAE symptoms and Karnofsky Performance Status longitudinally at sequential office visits. To compare how patient's vs clinician's reports relate to sentinel clinical events, a time-dependent Cox regression model was used to measure associations between reaching particular CTCAE grade severity thresholds with the risk of death and emergency room visits. To measure concordance of CTCAE reports with indices of daily health status, Kendall tau rank correlation coefficients were calculated for each symptom with EuroQoL EQ-5D questionnaire and global question scores. Statistical tests were two-sided.
- Results** A total of 163 patients were enrolled for an average of 12 months (range = 1–28 months), with a mean of 11 visits and 67 (41%) deaths. CTCAE reports were submitted by clinicians at 95% of visits and by patients at 80% of visits. Patients generally reported symptoms earlier and more frequently than clinicians. Statistically significant associations with death and emergency room admissions were seen for clinician reports of fatigue ($P < .001$), nausea ($P = .01$), constipation ($P = .038$), and Karnofsky Performance Status ($P < .001$) but not for patient reports of these items. Higher concordance with EuroQoL EQ-5D questionnaire and global question scores was observed for patient-reported symptoms than for clinician-reported symptoms.
- Conclusions** Longitudinally collected clinician CTCAE assessments better predict unfavorable clinical events, whereas patient reports better reflect daily health status. These perspectives are complementary, each providing clinically meaningful information. Inclusion of both types of data in treatment trial results and drug labels appears to be warranted.

J Natl Cancer Inst 2009;101:1624–1632

In cancer treatment trials, the standard source of adverse symptom data is clinician reporting by use of items from the National Cancer Institute's (NCI) Common Terminology Criteria for Adverse Events (CTCAE) (1). Patient self-reporting is not currently an accepted source of this information (2,3). Consequently, during the drug approval process, industry sponsors and the Food and Drug Administration (FDA) have access only to clinician impressions of adverse symptoms but not to patient accounts of these events (4). Toxicity tables in oncology drug labels include clinician-based information only. This approach has been criticized because of its exclusion of the patient's perspective, with suggestions that patients are in the best position to describe their own symptoms and that their perspective is relevant for labeling to inform future users of treatments about anticipated effects (5–7).

Recently, interest has increased at the FDA and NCI to adopt patient-reported outcomes in trials as a standard data source for measuring subjective phenomena (8–11). In 2006, a draft FDA Guidance was issued outlining standards for development of

patient-reported outcome measures to support labeling claims (12), followed by an NCI-sponsored conference on patient-reported outcomes in cancer research (13,14).

NCI is currently considering including a set of patient items in future versions of the CTCAE (2,15). The feasibility of collecting such data directly from patients has previously been demonstrated,

Affiliations of authors: Department of Medicine, Department of Nursing, and Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY (EB, XJ, GH, AB, LS, MF, MA, AI, TA, SG, AC, MGK); Dana-Farber/Harvard Cancer Center, Central Administration, 44 Binney Street, Boston, MA (DS).

Correspondence to: Ethan Basch, MD, MSc, Health Outcomes Group, Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, 307 East 63rd St, New York, NY 10065 (e-mail: basche@mskcc.org).

See "Funding" and "Notes" following "References."

DOI: 10.1093/jnci/djp386

© The Author 2009. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oxfordjournals.org.

Advance Access publication on November 17, 2009.

as has the willingness of practitioners to use this information to inform treatment decisions (16–18). Although the narrowing digital divide is likely to make electronic collection of such data more efficient and affordable in the future (19,20), widespread adoption of patient reporting would require substantial new resources at study sites (5) and would ultimately change how toxic drugs appear in their labels because patients receiving cancer treatment more often than clinicians report that symptoms are more severe (6,21,22).

Whether patient or clinician reports of adverse symptoms—or both—should be included in treatment trials and drug labels is a topic of recent debate (2,5). It is germane to this debate to determine which data source is more informative about the overall toxic impact of medical products on the patient experience because adverse event information is generally considered in aggregate during regulatory review as an indicator of safety (23). To make such a determination, we designed a study to compare the associations of longitudinally collected patient vs clinician adverse symptom reports by use of three established clinical outcomes that are central to the patient experience with cancer care: global health status, risk of emergency room visits, and risk of death.

Patients and Methods

Patients

A study protocol was approved by the Institutional Review Board of Memorial Sloan-Kettering Cancer Center. Patients with lung cancer starting noninvestigational chemotherapy regimens at Memorial Sloan-Kettering Cancer Center who were able to read and understand an English-language questionnaire were invited to enroll. All participants provided written informed consent. Patients were observed for up to 28 months or until death.

Study Design

At baseline and each clinic visit, patients were asked to complete a seven-item toxicity questionnaire via a touch screen tablet computer interface. Items included patient adaptations of CTCAE version 3.0 symptom items salient to individuals receiving chemotherapy (including fatigue, pain, nausea, vomiting, diarrhea, and constipation) and Karnofsky Performance Status (KPS). These items were previously developed for patient use through focus groups, interviews, and comparisons of patient vs clinician responses and have been previously described (16,17,21). Because these items were derived from the CTCAE, their intention by design is adverse symptom screening. Patients completed items before their encounters with clinicians because previous work has demonstrated no effects on either patient or clinician toxicity reporting, regardless of whether patients report before or after seeing clinicians (21). No incentives were offered for completing questionnaires. The study was conducted before the development of CTCAE version 4.0.

Clinician assessments of these same symptoms using the CTCAE, as well as KPS, are a part of standard chart documentation at Memorial Sloan-Kettering Cancer Center. At each clinical encounter, this information is completed by a nurse and/or a medical oncologist using a preprinted form. These data were abstracted from medical charts of enrollees for visit days during the study by

CONTEXT AND CAVEATS

Prior knowledge

Clinicians report adverse events in clinical trials using the National Cancer Institute's Common Terminology Criteria for Adverse Events; whether patient self-reports are useful is unknown.

Study design

Lung cancer patients who were receiving chemotherapy and their clinicians independently reported six Common Terminology Criteria for Adverse Events symptoms and Karnofsky Performance Status at office visits. Patient and clinician reports were compared by assessing associations between Common Terminology Criteria for Adverse Events grade severity thresholds and risk of death and emergency room visits. Relationships of Common Terminology Criteria for Adverse Events reports with indices of daily health status were also measured.

Contributions

Patients reported symptoms earlier and more frequently than clinicians. Associations with death and emergency room admissions were seen for clinician reports of fatigue, nausea, constipation, and performance status but not for patient reports of these items. Patient-reported symptoms were more often in agreement with measures of daily health status than clinician-reported symptoms.

Implications

The perspectives of clinicians and patients regarding adverse events are complementary—clinician reports better predict unfavorable clinical events and patient reports better reflect daily health status.

Limitations

Because the study included patients from one center with one cancer type and specific survey questions, the findings may not apply to other patient populations.

From the Editors

two trained research assistants (A. Barz and M. Appawu). It was documented whether the data source was a nurse or a physician, and if both reported independently at a given visit, each score was coded separately for use in a sensitivity analysis. In cases of coding disagreement or unclear reporting, the chart was reviewed by a senior data manager and medical oncologist for adjudication.

CTCAE items were graded for severity by use of a five-point ordinal scale, whereas an 11-point scale was used for KPS to assess level of physical capability (assigned in units of 10 between 0 and 100) and for pain as noted in the medical chart (assigned between 0 and 10). Per NCI instructions for the CTCAE, a score of 1 was defined as “mild,” 2 as “moderate,” 3 as “severe,” and 4 as “disabling” (1). For the 0–10 pain scale, previously established score ranges associated with mild, moderate, or severe designations were used (24–26).

Statistical Analysis

To compare the extent to which patient-reported vs clinician-reported items, which had been collected at multiple time points longitudinally, predicted discrete clinical outcomes, a time-dependent Cox model was used in which each possible grade level was considered as a threshold for each item of interest (ie, mild, moderate, severe, or

Table 1. Characteristics of patients

Characteristic	No. of patients (n = 163)	%
Age		
Mean, y		63
Median, y		63
18–39	4	2
40–49	16	10
50–59	39	24
60–69	59	36
>70	45	28
Sex		
Female	89	55
Cancer type		
Non–small cell lung cancer	136	83
Localized (stages I, II, and IIIA)	42	31
Advanced/metastatic (stages IIIB and IV)	94	69
Small cell lung cancer	15	9
Limited stage	6	40
Extensive stage	9	60
Thymoma	6	4
Mesothelioma	6	4
Baseline ECOG score*		
0	52	32
1	105	64
2	6	4
Baseline subjective health stat†		
90–100	39	25
80–89	33	22
70–79	36	24
60–69	16	10
50–59	15	10
≤49	14	10
Highest education level		
Professional/graduate degree	35	21
College degree	40	25
Some college	30	18
High school or less	58	36

* Staff-reported, Eastern Cooperative Oncology Group performance status criteria.

† Patient-reported global health status from 0–100 via a validated single-item visual analog scale (28).

disabling). Clinical outcomes included death and emergency room visits. The association of each specified grade threshold with the risk of each clinical outcome was calculated. This analysis was conducted separately for the patient-reported version and clinician-reported version of each item to compare their relative strengths of association. *P* values were generated with Wald statistics. Only those predictor values that were recorded before a clinical outcome of interest were used in the Cox model.

To compare the extent to which patient-reported vs clinician-reported items reflect global health status, patients completed two measures during the study: 1) the EuroQoL EQ-5D questionnaire (27) and 2) a global question that uses a visual analog scale to rate current health state from worst to best imaginable (28). Both measures have been shown to be reliable and valid assessments of health status in cancer patients (28,29). The EQ-5D consists of five items (including mobility, self-care, daily activities, pain and/or discomfort, and anxiety and/or depression), which are combined to render a single score that is adjusted for US population prefer-

ence weights. The global question yields a single score between 0 and 100. Patients completed the EQ-5D questionnaire at each clinic visit and the global question only at baseline. To compare the strength of concordance of each patient-reported and clinician-reported item with these two measures of health status, Kendall tau rank correlation coefficients were calculated for each relationship, with -1 representing perfect negative concordance, 0 representing independence, and $+1$ representing perfect positive concordance. Confidence intervals (CIs) were derived from bootstrap resamplings on patients to account for correlations among repeated measures.

For each successive visit, the proportion of symptom reports completed by patients vs clinicians was tabulated to compare the extent of missing data between these two reporting approaches. If a patient’s or a clinician’s report was missing at a visit, bracketed by completed reports at previous and subsequent visits, then the most recent previously reported score was used for the analysis.

P values were generated by use of the two-sided Wald test for parameter estimates in the Cox models; two-sided *P* values and 95% confidence intervals for the Kendall tau correlation coefficients were generated by bootstrap resamplings on patients to account for correlations among repeated measurements. *P* values of less than .05 were considered as statistically significant.

Results

Between June 15, 2005, and June 14, 2006, 190 consecutive patients with lung cancer were approached, of whom 185 were eligible and 163 consented to participate. Baseline characteristics are shown in Table 1. The median age was 63 years (range = 35–85 years). Most patients (69%) were diagnosed with advanced or metastatic non–small cell lung cancer, with predominantly good baseline provider-reported performance status. A majority of patients received cytotoxic chemotherapy in clinic once every 3 weeks. Mean enrollment was 12 months (range = 1–28 months), and the average number of clinic visits was 11 (range = 1–40). Sixty-seven (41%) of the 163 patients died during the study period. Twenty clinicians cared for these patients and documented their adverse symptoms in medical charts, including nine attending oncologists and 11 oncology nurses. The average number of patients cared for by any given oncologist was 16, with three oncologists treating fewer than 10 patients, three treating 10–20, and three treating 30–40. During the study, there were a total of 1712 clinic visits at which adverse symptom forms were submitted by patients at 1362 visits (80%) and by clinicians at 1601 visits (95%). Rates of missing data varied by symptom (Table 2). In general, rates of survey completion were higher among clinicians than patients. Reasons cited by patients for not completing forms included “no changes in symptoms from the previous time reporting” (56%), “inconvenient timing” (15%), “staff forgot to administer form” (14%), “felt too sick” (8%), and “did not feel like it” (6%). There was no effect of treatment type on compliance levels. At most visits, clinician documentation of adverse symptoms was completed by nurses (53%), followed by attending oncologists (38%), and oncology fellows (7%).

Figure 1 shows the cumulative incidence of attaining various severity levels of symptom items and performance status as

Table 2. Proportion of patients and clinicians completing reports for each symptom at successive clinic visits

Report source	% Completing report											
	Visit 1 (n = 163)	Visit 2 (n = 156)	Visit 3 (n = 148)	Visit 4 (n = 137)	Visit 5 (n = 123)	Visit 6 (n = 108)	Visit 7 (n = 97)	Visit 8 (n = 93)	Visit 9 (n = 80)	Visit 10 (n = 72)	Visit 11 (n = 67)	Visit 12 (n = 62)
Patient												
Fatigue	93	84	83	82	76	79	78	75	76	72	75	79
Pain	93	83	82	79	74	77	75	73	75	71	75	77
Nausea	93	83	82	81	75	78	78	75	75	72	75	79
Vomiting	92	83	81	82	76	75	76	73	76	72	73	77
Diarrhea	93	83	82	81	75	76	78	75	74	72	75	79
Constipation	93	84	82	82	75	76	78	73	76	72	75	79
Clinician												
Fatigue	98	93	95	91	95	89	89	92	88	89	88	89
Pain	83	80	76	80	84	75	70	71	75	74	67	73
Nausea	98	97	95	95	95	94	91	94	94	92	87	90
Vomiting	100	96	95	94	95	93	92	95	94	94	88	89
Diarrhea	99	96	93	92	94	92	91	91	93	92	88	87
Constipation	98	94	94	90	95	91	89	94	91	89	87	85

reported by patients vs clinicians, with death considered as a competing risk. In general, each symptom severity level (ie, threshold) was reached earlier in the course of care and more frequently according to patient reports as opposed to clinician reports.

Table 3 shows the associations of symptoms reported by patients or clinicians with the risk of death and the risk of emergency room visits. For the threshold of moderate fatigue in a univariate analysis, there was no relationship between patient self-reports and the risk of death ($P = .23$), whereas there was a highly statistically significant relationship for clinician reports of this symptom severity level ($P < .001$), with a hazard ratio for death of 2.75. The threshold was reached by a similar number of patients by each reporting approach, with a similar number of deaths among those reaching the threshold (46% and 48%). Raising the threshold to severe fatigue yielded similar results, with a non-statistically significant relationship for patient self-reports ($P = .15$), but a statistically significant relationship for clinician reports ($P = .003$), with a hazard ratio for death of 2.39. In this case, fewer patients reached the threshold by the clinician's report compared with the patient's report (37% vs 17%), yet a relatively higher proportion of those identified by clinicians as having severe fatigue died (57% vs 48%).

A similar pattern was observed for the relationships of moderate nausea and constipation thresholds with the risk of death, with non-statistically significant relationships for patient reports ($P = .38$ and $P = .40$, respectively) and statistically significant relationships for clinician reports ($P = .01$ and $P = .038$, respectively). Too few reports of severe nausea or vomiting existed to conduct analyses of these thresholds. The pattern was repeated for KPS, for which clinician reports yielded the strongest observed association (hazard ratio = 6.39, 95% CI 3.88 to 11.08; $P < .001$) compared with patient reports (hazard ratio = 1.40, 95% CI = 0.84 to 2.35; $P = .20$). The cutoff KPS score of 70 or less [defined as "cares for self; unable to carry on normal activity or to do active work" (30)] was the most predictive threshold, although other cutoffs yielded similar results. Notably, worse survival was seen in patients for whom clinicians reported severe nausea, pain, fatigue, or KPS score of 70 or less during the initial month of observation (with a

median survival time of 12 months) compared with patients for whom severe grades were not assigned (who did not reach a median survival time during the study).

Analyses were generally not informative for symptoms when using mild as a threshold (because of commonness at baseline by both patient and clinician reports) or when using disabling as a threshold (because of its overall rarity in this outpatient population). Too few events existed for any threshold levels of vomiting or diarrhea to conduct analyses of these symptoms. No effect on any estimates by patient characteristics including sex, age, education level, clinician type (physician or nurse), treating oncologist, disease stage, or chemotherapy type was observed.

A sensitivity analysis evaluated the associations between the second time patients reached each severity threshold and the outcomes of interest, and results were similar. In a bivariate analysis, patient-reported data did not augment the predictive ability of clinician reporting. In a multivariable model that included all items, KPS dominated in its ability to predict the measured endpoints.

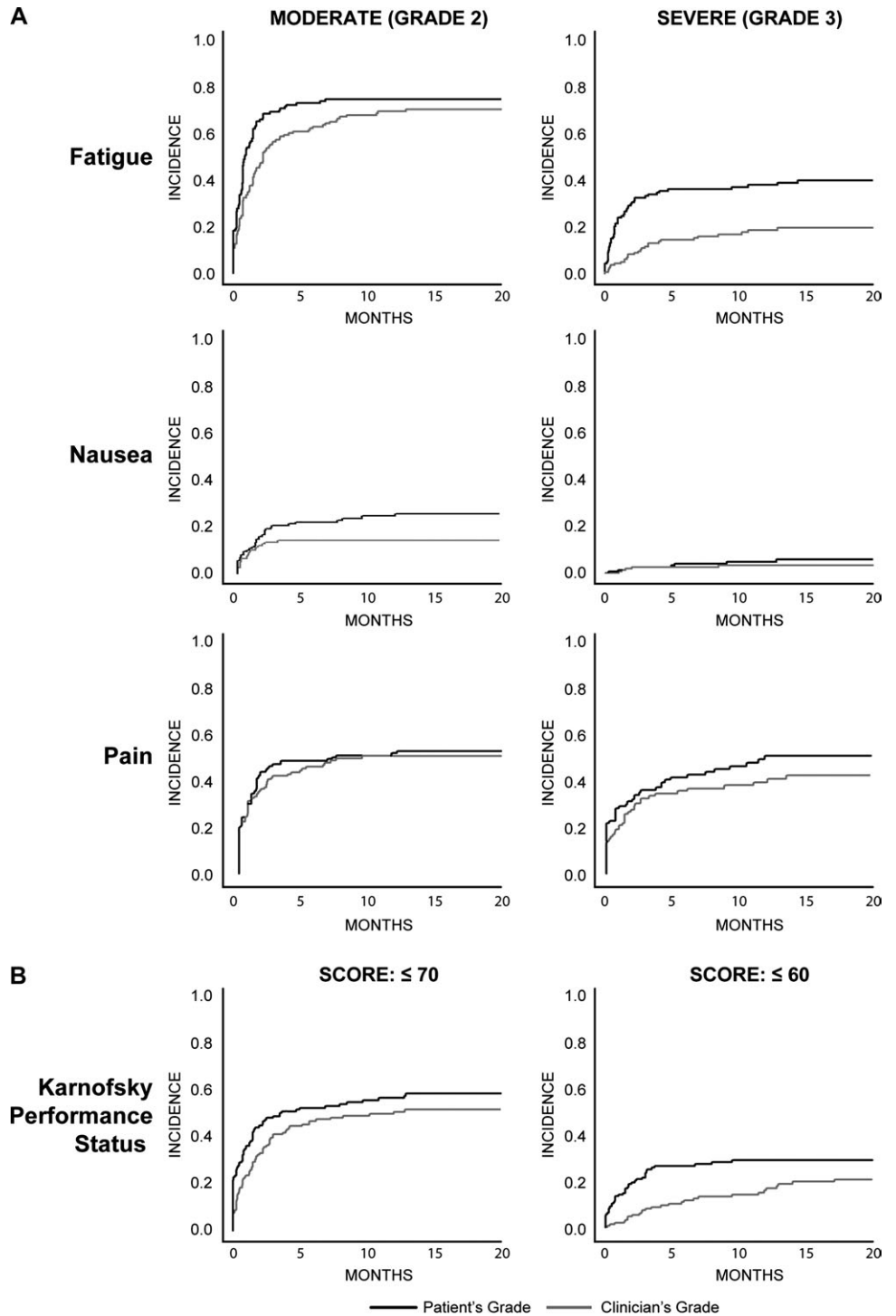
Results were similar for the association of symptom thresholds with risk of emergency room admission, with statistically significant relationships observed for clinician-reported moderate fatigue, pain, and KPS thresholds but not for patient reports. There were too few events for other thresholds to conduct analyses.

Figure 2 shows the relative strengths of concordance between patient-reported and clinician-reported CTCAE symptoms and KPS with the two validated measures of health status (EuroQoL EQ-5D questionnaire and 0–100 global question). In general, higher levels of concordance were seen for patient reports compared with clinician reports. Results were more pronounced for EQ-5D scores, which were measured at each sequential clinic visit, compared with 0–100 global question scores, which were measured only at baseline.

Discussion

In this longitudinal study of patients with advanced lung cancer, we found that clinician reporting of CTCAE symptoms and performance status was statistically significantly associated with

Figure 1. Cumulative incidence of symptoms and performance status as reported by patients vs clinicians and by month of follow-up, with death considered as a competing risk. **A)** Moderate (grade 2) or severe (grade 3) symptoms according to the National Cancer Institute’s Common Terminology Criteria for Adverse Events, version 3.0. **B)** Karnofsky Performance Status. A Karnofsky score of 70 is defined as “Cares for self; unable to carry on normal activity or to do active work,” whereas a score of 60 is defined as “Requires occasional assistance, but is able to care for most of his personal needs” (30).



unfavorable clinical outcomes, such as death and emergency room admissions, whereas patient reporting was more strongly associated with measures of daily health status. Patients generally reported symptoms earlier and more frequently than clinicians. Because clinicians are attuned to patients’ trajectories to major disease benchmarks, they appeared to reserve assignment of more severe grades until it was clear that such a benchmark was impending at the expense of sensitivity to granular aspects of patients’ daily experiences. In contrast, patient reporting appeared to better reflect real-time suffering at the expense

of sensitivity to impending sentinel events, such as death or hospitalization. These results expand on previous studies that have reported statistically significant associations of clinician-reported performance status with survival (31–33) and of patient-reported outcomes including pain and health-related quality of life with survival (34–41).

Patient and clinician perspectives of adverse symptoms appear to be complementary, together providing a more complete picture of the toxic impact of treatments compared with either perspective alone. Clinicians bring professional training and experience to

Table 3. Associations of patient-reported or clinician-reported symptoms and performance status with risk of death or risk of ER admission (n = 163)*

Threshold	Reporter	Risk of death		Risk of ER admission		% Reaching threshold before death	% Dying who reached threshold
		HR (95% CI)	P	HR (95% CI)	P		
Fatigue							
Moderate	Patient	1.46 (0.79 to 2.69)	.23	1.18 (0.75 to 1.83)	.48	71	46
	Clinician	2.75 (1.54 to 4.91)	<.001	1.87 (1.23 to 2.86)	.004	66	48
Severe	Patient	1.44 (0.89 to 2.37)	.15	†	†	37	48
	Clinician	2.39 (1.43 to 4.46)	.003	†	†	17	57
Nausea†							
Moderate	Patient	1.27 (0.74 to 2.18)	.38	†	†	33	45
	Clinician	2.07 (1.18 to 3.63)	.01	†	†	16	65
Constipation†							
Moderate	Patient	1.26 (0.74 to 2.13)	.40	1.53 (0.95 to 2.45)	.08	28	47
	Clinician	1.68 (1.03 to 2.75)	.038	1.28 (0.82 to 2.00)	.28	39	48
KPS							
Score ≤80	Patient	1.48 (0.92 to 2.51)	.13	1.27 (0.84 to 1.910)	.26	55	46
	Clinician	4.91 (2.89 to 8.38)	<.001	1.68 (1.07 to 2.62)	.02	48	59
Score ≤70	Patient	1.40 (0.84 to 2.35)	.20	†	†	28	48
	Clinician	6.39 (3.88 to 11.08)	<.001	†	†	18	77
Pain							
Moderate	Patient	1.47 (0.89 to 2.41)	.13	1.44 (0.95 to 2.180)	.09	50	47
	Clinician	1.56 (0.94 to 2.59)	.08	1.84 (1.22 to 2.76)	.004	47	51
Severe	Patient	2.18 (1.22 to 3.89)	.008	1.71 (0.97 to 2.99)	.06	24	56
	Clinician	2.29 (1.39 to 3.78)	.001	1.67 (1.05 to 2.68)	.03	29	61

* CI = confidence interval; ER = emergency room; HR = hazard ratio; KPS = Karnofsky Performance Status. P values (two-sided) were generated by using the two-sided Wald test for parameter estimates.

† Too few deaths among patients reaching severe nausea or constipation thresholds, and too few ER admissions among patients reaching severe fatigue, moderate or severe nausea, severe constipation, or KPS 70 or less thresholds to conduct analyses.

their evaluations, whereas patients are in a better position to communicate their own subjective experiences. Currently, in cancer treatment trials, clinicians but not patients serve as the source of adverse symptom data (1). Our findings demonstrate that this approach resulted in the loss of information that might be valuable to prospective prescribers and patients in understanding the anticipated effects of treatment.

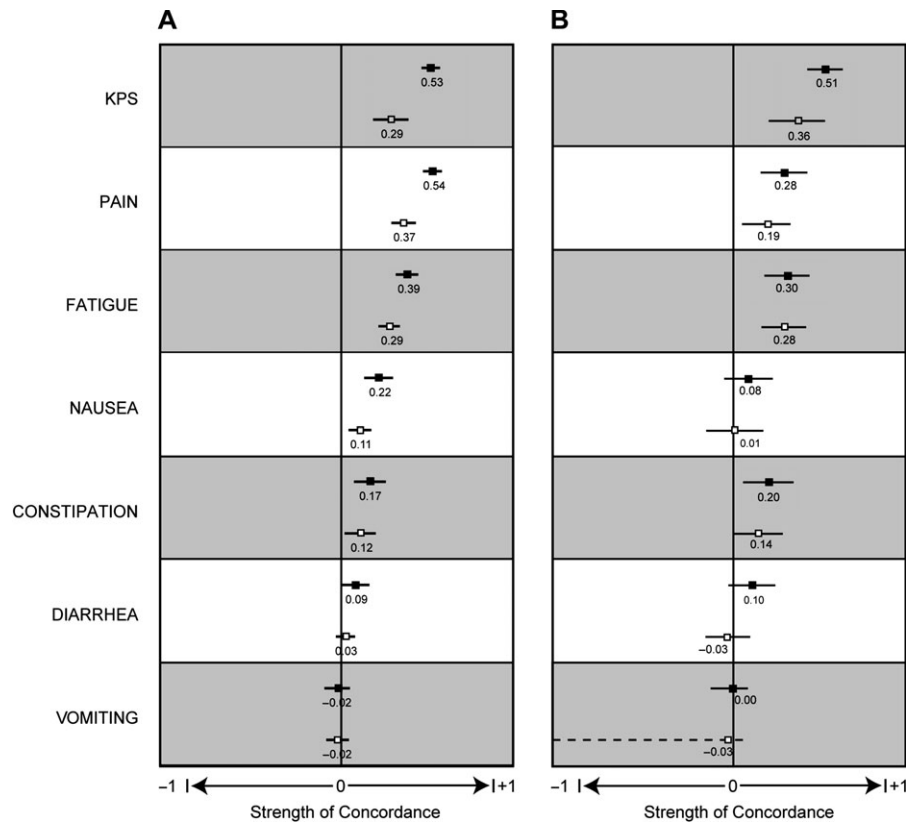
As this study and previous work (6,21,22) have demonstrated, patients more frequently report worse symptom severity than clinicians. We further found that patients tend to report adverse symptoms earlier in the course of care than their clinicians. Therefore, drug labels that are based exclusively on clinician reporting invariably underestimate the frequency and severity of adverse symptom events as compared with the patient's perspective. In contrast, if patient reporting were adopted as a new standard, then a frameshift would occur whereby drugs would appear more toxic in their labels. This frameshift effect would balance out in controlled studies because symptoms unrelated to treatment would presumably equilibrate between groups. Because the incidence of reported symptoms would rise across the entire patient population, such trials would have greater power to discern between background symptoms and those related to interventions. In single-arm trials, comparisons of patient-reported adverse symptoms with baseline values would help distinguish between preexisting symptoms and those that developed during a study. Because single-arm trials are uncommonly used as data sources for cataloging adverse events in labels, symptom reports from phase III trials would have the advantage of identifying the incremental symptom burden attributable to new treatments.

Given our findings, which source of adverse symptom data should be included in oncology trials and drug labels: clinicians, patients, or both? Because clinical trials already capture survival and hospitalization data, and clinician-reported symptoms appear to be closely related to these events but not with daily health status, we suggest that clinician-reported symptoms will add only limited information beyond what is already collected in trials. In contrast, patient reports are less frequently associated with clinical endpoints but do reflect daily health status and, hence, provide additional information beyond what is currently collected in trials or reported in labels. The addition of patient-reported adverse symptoms to drug labels may therefore be warranted.

Even if patient reporting of adverse symptoms is adopted in oncology trials, it is unlikely that clinician reporting of this information will be discontinued. Clinician reporting is well established and, as demonstrated in this study, is generally more complete than patient reporting. Therefore, the questions are whether the current model should be expanded to include documentation of patient-reported outcomes in addition to clinician reports and whether it would cause confusion if clinician and patient adverse symptom data were both presented in study results and labels, given that they are often discrepant (6,21,22) (eg, would a label confuse readers if 30% of clinicians but 55% of patients reported severe nausea?).

In fact, there are many examples of reporting structures outside of medicine in which professional and nonprofessional ratings of the same phenomena are presented together and are well accepted. Highly trafficked Internet review sites for books, hotels, movie reviews, and consumer electronics provide expert and consumer

Figure 2. Relative strengths of concordance of patient-reported and clinician-reported Karnofsky Performance Status and symptoms with two measures of health status (28,29). **A)** Concordance with EuroQoL EQ-5D questionnaire scores. **B)** Concordance with 0–100 global question scores. **Solid squares** = patient reported; **open squares** = clinician reported. For each relationship, Kendall tau rank correlation coefficients are shown with 95% confidence intervals, with -1 representing perfect negative concordance, 0 representing independence, and $+1$ representing perfect positive concordance.



ratings side-by-side, despite frequent discrepancies (19). Implicit in these sites is an understanding that professionals and nonprofessionals consider different criteria when reviewing the same product, and each perspective brings value. When considering a hotel, the opinion of an expert with a broad and technical standpoint is useful, as are the views of individuals who have actually stayed at the hotel. However, drug labels are not like these Web sites and, by comparison, are missing half the picture. Consequently, when an oncologist and patient sit down to discuss starting a new regimen and wish to consider its potential toxic effects, the impressions of that patient's peers are not available—only the impressions of the oncologist's colleagues are included in the label.

Table 4. Potential beneficiaries of patient-reporting of adverse symptoms in cancer treatment trials*

Stakeholder	Potential benefits
Clinical trial enrollees	Earlier detection of toxic effects through improved communication with clinical staff
Investigators and/or sponsors	More complete adverse event data during drug development
FDA reviewers	Additional toxicity data to balance safety with efficacy during regulatory review
Clinicians	Improved information about prior patients' experiences with treatments, for use when counseling future patients or assessing adverse reactions
Future patients	Access to information about prior patients' experiences with particular treatments, to inform therapy decisions

* FDA = US Food and Drug Administration.

The value of patient input in clinical research is already well established in other contexts, including patient advocate participation in NCI study sections and FDA advisory committee meetings and patient involvement in trials directly reporting health-related quality of life and symptom endpoint data. Results of this study indicate the direction that extending this model to incorporate the patient's perspective in adverse event documentation could take. Such a change would potentially have implications for multiple stakeholders, including patients, investigators, FDA reviewers, industry sponsors, and clinicians (Table 4) (18,42). In the future, this patient-reporting model could be expanded into the routine care setting to assist clinicians with symptom management and could also contribute real-time data to postmarket drug safety surveillance systems (43,44).

This study had several limitations. It was conducted at an urban tertiary cancer center in patients with a single cancer type, potentially limiting the generalizability of our findings. The results are based on the use of specific symptom and performance status items and may be subject to unique measurement properties that may not extend to other items, although in general, the items used are standard for assessing patient status in NCI-sponsored treatment trials (1). Corroboration of the results is warranted in a multicenter study with a diverse patient population, and such a study is under way in the NCI cooperative group setting. Nonresponse bias is a common challenge in symptom research, although response rates were high overall in this study and reasons for nonresponses were tracked. It is not known whether patients and their clinicians would alter treatment decisions when provided with patient-reported data about drug toxicity. Future studies in this area would provide an indication

of the extent to which patterns of care would change if drug labels included patient-reported adverse symptom information.

References

1. Cancer Therapy Evaluation Program NCI. *Common Terminology Criteria for Adverse Events (CTCAE) v3.0 Online Instructions and Guidelines*. https://webapps.ctep.nci.nih.gov/webobjs/ctc/webhelp/welcome_to_ctcae.htm. Accessed May 19, 2009.
2. 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):5121–5127.
3. Gwede CK, Johnson DJ, Daniels SS, Trotti A. Assessment of toxicity in cooperative oncology clinical trials: the long and short of it. *J Oncol Manag*. 2002;11(2):15–21.
4. Code of Federal Regulations, Title 21, Volume 5. Part 312—Investigational New Drug Application, Section 312.32, IND safety reports. Revised April 1, 2008 (21CFR312.32).
5. Bruner DW. Should patient-reported outcomes be mandatory for toxicity reporting in cancer clinical trials? *J Clin Oncol*. 2007;25(34):5345–5347.
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):3485–3490.
7. Brundage MD, Pater JL, Zee B. Assessing the reliability of two toxicity scales: implications for interpreting toxicity data. *J Natl Cancer Inst*. 1993;85(14):1138–1148.
8. Acquadro C, Berzon R, Dubois D, et al. Incorporating the patient's perspective into drug development and communication: an ad hoc task force report of the Patient-Reported Outcomes (PRO) Harmonization Group meeting at the Food and Drug Administration, February 16, 2001. *Value Health*. 2003;6(5):522–531.
9. Morris LA, Miller DW. The regulation of patient-reported outcome claims: need for a flexible standard. *Value Health*. 2002;5(4):372–381.
10. Rock EP, Kennedy DL, Furness MH, Pierce WF, Pazdur R, Burke LB. Patient-reported outcomes supporting anticancer product approvals. *J Clin Oncol*. 2007;25(32):5094–5099.
11. Lipscomb J, Reeve BB, Clauser SB, et al. Patient-reported outcomes assessment in cancer trials: taking stock, moving forward. *J Clin Oncol*. 2007;25(32):5133–5140.
12. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006 Oct 11;4:79.
13. Clauser SB, Ganz PA, Lipscomb J, Reeve BB. Patient-reported outcomes assessment in cancer trials: evaluating and enhancing the payoff to decision making. *J Clin Oncol*. 2007;25(32):5049–5050.
14. Ganz PA, Gotay CC. Use of patient-reported outcomes in phase III cancer treatment trials: lessons learned and future directions. *J Clin Oncol*. 2007;25(32):5063–5069.
15. Office of Acquisitions, National Cancer Institute, National Institutes of Health, Department of Health and Human Services. Request for Proposals Solicitation Number N02-PC-85002–29: “Development of the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE).” https://www.fbo.gov/index?s=opportunity&mode=form&tab=core&id=42380ab82f5791a814ecb0d22a8db319&_cview=0. Accessed May 19, 2009.
16. Basch E, Artz D, Dulko D, et al. Patient online self-reporting of toxicity symptoms during chemotherapy. *J Clin Oncol*. 2005;23(15):3552–3561.
17. Basch E, Iasonos A, Barz A, et al. Long-term toxicity monitoring via electronic patient-reported outcomes in patients receiving chemotherapy. *J Clin Oncol*. 2007;25(34):5374–5380.
18. Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol*. 2004;22(4):714–724.
19. Alexa, The Web Information Company. Traffic Rankings, Top Sites United States. http://www.alexa.com/site/ds/top_sites?cc=US&ts_mode=country&lang=none. Accessed May 19, 2009.
20. Nielsen//NetRatings. U.S. broadband composition reaches 72 percent at home, a 15 point year-over-year increase. http://www.nielsen-online.com/pr/pr_060621.pdf. Accessed May 19, 2009.
21. Basch E, Iasonos A, McDonough T, et al. 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):903–909.
22. Vistad I, Cvancarova M, Fossa SD, Kristensen GB. Postradiotherapy morbidity in long-term survivors after locally advanced cervical cancer: how well do physicians' assessments agree with those of their patients? *Int J Radiat Oncol Biol Phys*. 2008;71(5):1335–1342.
23. U.S. Department of Health and Human Services, Food and Drug Administration. Guidance for industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics. Released May 2007.
24. Chow E, Doyle M, Li K, et al. Mild, moderate, or severe pain categorized by patients with cancer with bone metastases. *J Palliat Med*. 2006;9(4):850–854.
25. Palos GR, Mendoza TR, Mobley GM, Cantor SB, Cleeland CS. Asking the community about cutpoints used to describe mild, moderate, and severe pain. *J Pain*. 2006;7(1):49–56.
26. Serlin RC, Mendoza TR, Nakamura Y, Edwards KR, Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. *Pain*. 1995;61(2):277–284.
27. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med*. 2001;33(5):337–343.
28. de Boer AG, van Lanschot JJ, Stalmeier PF, et al. Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? *Qual Life Res*. 2004;13(2):311–320.
29. Brazier J, Jones N, Kind P. Testing the validity of the Euroqol and comparing it with the SF-36 health survey questionnaire. *Qual Life Res*. 1993;2(3):169–180.
30. Yates JW, Chalmer B, McKeegney FP. Evaluation of patients with advanced cancer using the Karnofsky performance status. *Cancer*. 1980;45(8):2220–2224.
31. Chau I, Norman AR, Cunningham D, Waters JS, Oates J, Ross PJ. Multivariate prognostic factor analysis in locally advanced and metastatic esophago-gastric cancer—pooled analysis from three multicenter, randomized, controlled trials using individual patient data. *J Clin Oncol*. 2004;22(12):2395–2403.
32. Hauser CA, Stockler MR, Tattersall MH. Prognostic factors in patients with recently diagnosed incurable cancer: a systematic review. *Support Care Cancer*. 2006;14(10):999–1011.
33. Loprinzi CL, Laurie JA, Wieand HS, et al. Prospective evaluation of prognostic variables from patient-completed questionnaires. North Central Cancer Treatment Group. *J Clin Oncol*. 1994;12(3):601–607.
34. Gotay CC, Kawamoto CT, Bottomley A, Efficace F. The prognostic significance of patient-reported outcomes in cancer clinical trials. *J Clin Oncol*. 2008;26(8):1355–1363.
35. Halabi S, Vogelzang NJ, Kornblith AB, et al. Pain predicts overall survival in men with metastatic castration-refractory prostate cancer. *J Clin Oncol*. 2008;26(15):2544–2549.
36. Coates AS, Hurny C, Peterson HF, et al. Quality-of-life scores predict outcome in metastatic but not early breast cancer. International Breast Cancer Study Group. *J Clin Oncol*. 2000;18(22):3768–3774.
37. Collette L, van Andel G, Bottomley A, et al. Is baseline quality of life useful for predicting survival with hormone-refractory prostate cancer? A pooled analysis of three studies of the European Organisation for Research and Treatment of Cancer Genitourinary Group. *J Clin Oncol*. 2004;22(19):3877–3885.
38. Efficace F, Bottomley A, Smit EF, et al. Is a patient's self-reported health-related quality of life a prognostic factor for survival in non-small-cell lung cancer patients? A multivariate analysis of prognostic factors of EORTC study 08975. *Ann Oncol*. 2006;17(11):1698–1704.
39. Ganz PA, Lee JJ, Siau J. Quality of life assessment. An independent prognostic variable for survival in lung cancer. *Cancer*. 1991;67(12):3131–3135.

40. Langendijk H, Aaronson NK, de Jong JM, ten Velde GP, Muller MJ, Wouters M. The prognostic impact of quality of life assessed with the EORTC QLQ-C30 in inoperable non-small cell lung carcinoma treated with radiotherapy. *Radiother Oncol.* 2000;55(1):19–25.
41. Yeo W, Mo FK, Koh J, et al. Quality of life is predictive of survival in patients with unresectable hepatocellular carcinoma. *Ann Oncol.* 2006;17(7):1083–1089.
42. Detmar SB, Muller MJ, Schornagel JH, Wever LD, Aaronson NK. Health-related quality-of-life assessments and patient-physician communication: a randomized controlled trial. *JAMA.* 2002;288(23):3027–3034.
43. Bennett CL, Nebeker JR, Yarnold PR, et al. Evaluation of serious adverse drug reactions: a proactive pharmacovigilance program (RADAR) vs safety activities conducted by the Food and Drug Administration and pharmaceutical manufacturers. *Arch Intern Med.* 2007;167(10):1041–1049.
44. Kuehn BM. FDA turns to electronic “sentinel” to flag prescription drug safety problems. *JAMA.* 2008;300(2):156–157.

Funding

Funding to conduct this research was received from the National Cancer Institute, American Society of Clinical Oncology, and the Steps-for-Breath philanthropic fund of Memorial Sloan-Kettering Cancer Center.

Notes

The sponsors had no role in the study design, data collection or analysis, interpretation of the results, or in the preparation of the manuscript or in the decision to submit the manuscript for publication.

The authors wish to thank Drs Elena Elkin and Yuelin Li for reviewing the manuscript and providing valuable comments, and Tony Riley and Joseph Christoff for their assistance with graphics and document preparation, respectively.

Manuscript received February 3, 2009; revised August 28, 2009; accepted September 28, 2009.