Symptomatic Toxicities Experienced During Anticancer Treatment: Agreement Between Patient and Physician Reporting in Three Randomized Trials

Massimo Di Maio, Ciro Gallo, Natasha B. Leight, Maria Carmela Piccirillo, Gennaro Daniele, Francesco Nuzzo, Cesare Gridelli, Vittorio Gębba, Fortunato Cardiello, Sabino De Placido, Anna Ceribelli, Adolfo G. Favaretto, Andrea de Matteis, Ronald Feld, Charles Butts, Jane Bryce, Simona Signoriello, Alessandro Morabito, Gaetano Rocco, and Francesco Perrone

Listen to the podcast by Dr Snyder at www.jco.org/podcasts

ABSTRACT

Purpose Information about symptomatic toxicities of anticancer treatments is not based on direct report by patients, but rather on reports by clinicians in trials. Given the potential for under-reporting, our aim was to compare reporting by patients and physicians of six toxicities (anorexia, nausea, vomiting, constipation, diarrhea, and hair loss) within three randomized trials.

Patients and Methods In one trial, elderly patients with breast cancer received adjuvant chemotherapy; in two trials, patients with advanced non–small-cell lung cancer received first-line treatment. Toxicity was prospectively collected by investigators (graded by National Cancer Institute Common Toxicity Criteria [version 2.0] or Common Terminology Criteria for Adverse Events [version 3]). At the end of each cycle, patients completed the European Organisation for Research and Treatment of Cancer quality-of-life questionnaires, including toxicity-related symptom items. Possible answers were “not at all,” “a little,” “quite a bit,” and “very much.” Analysis was limited to the first three cycles. For each toxicity, agreement between patients and physicians and under-reporting by physicians (ie, toxicity reported by patients but not reported by physicians) were calculated.

Results Overall, 1,090 patients (2,482 cycles) were included. Agreement between patients and physicians was low for all toxicities. Toxicity rates reported by physicians were always lower than those reported by patients. For patients who reported toxicity (any severity), under-reporting by physicians ranged from 40.7% to 74.4%. Examining only patients who reported “very much” toxicity, under-reporting by physicians ranged from 13.0% to 50.0%.

Conclusion Subjective toxicities are at high risk of under-reporting by physicians, even when prospectively collected within randomized trials. This strongly supports the incorporation of patient-reported outcomes into toxicity reporting in clinical trials.

J Clin Oncol 33:910-915. © 2015 by American Society of Clinical Oncology

INTRODUCTION

Information available to oncologists and their patients about symptomatic toxicities of anticancer treatments is not based on direct report by prior patients, but instead on reports made by clinician assessment in clinical trials.1 The potential for under-reporting of toxicity does not necessarily bias the direct comparison between the treatments, if the under-reporting rate is similar between treatment arms. However, it could substantially affect absolute estimates of toxicity, which is highly relevant for the applicability of trial results in clinical practice, particularly for new drugs.

Approximately 20 years ago, the Outcomes Working Group of the American Society of Clinical Oncology Health Research Committee defined the outcomes to be used for technology assessment and development of cancer treatment guidelines. Toxicity was considered a vitally important outcome.4 In
that article, the authors underlined that the evaluation of a subjective toxicity starts by explicitly asking the patient whether that toxicity occurred, highlighting that even if the patient does not report toxic symptoms voluntarily, this does not imply that the toxicity did not occur.

In a modern view of evidence-based practice, satisfying communication between patients with cancer and their physicians about benefits and risks associated with treatments is a critical component of care, ensuring that patients’ preferences are taken into account in decision making. In recent years, reflecting an increasing focus on a patient-centered approach, scientific interest in the integration of patient-reported outcomes into drug safety evaluation and comparative-effectiveness research is growing.

The aim of this study was to describe patients’ and physicians’ reporting of six symptomatic toxicities occurring during anticancer treatment, based on data prospectively collected in randomized trials, to describe the agreement between patients’ and physicians’ reports and the rate of possible under-reporting by physicians.

Patients enrolled onto three randomized controlled trials, all coordinated by the Clinical Trials Unit of the National Cancer Institute (Napoli, Italy) were included in this analysis. In the ELDA (Elderly Breast Cancer—Docetaxel Adjuvant) trial (ClinicalTrials.gov identifier NCT00331097), patients age 65 years with advanced NSCLC and ECOG performance status of 0 to 1 were randomly assigned to first-line cisplatin 80 mg/m² IV on days 1 and 8 plus docetaxel 35 mg/m² on days 1, 8, and 15 every 4 weeks. Between July 2003 and May 2005, 400 patients were enrolled in Italian institutions.

In the GECO (Gemcitabine–Cisdx) trial of non–small-cell lung cancer (NSCLC; ClinicalTrials.gov identifier NCT00338066), patients age < 70 years with advanced NSCLC and ECOG performance status of 0 to 1 were randomly assigned to one of four treatment groups: arm A, gemcitabine 1,200 mg/m² in a 30-minute IV infusion on days 1 and 8 plus cisplatin 80 mg/m² IV on day 1, every 21 days; arm B, same as arm A plus oral rofecoxib 50 mg daily; arm C, prolonged constant IV gemcitabine 1,200 mg/m² over 120-minute infusion on days 1 and 8 plus cisplatin 80 mg/m² IV on day 1, every 21 days; and arm D, same as arm C plus oral rofecoxib 50 mg daily. Between January 2003 and April 2011, 299 patients were randomly assigned in Italian institutions.

In the TORCH (Tarceva or Chemotherapy) trial (ClinicalTrials.gov identifier NCT00349219), patients with advanced NSCLC and ECOG performance status of 0 to 1 were randomly assigned to first-line cisplatin 80 mg/m² IV on day 1 plus gemcitabine 1,200 mg/m² IV on days 1 and 8 plus erlotinib 150 mg per day orally; erlotinib administration was conventionally divided into 3-week cycles. Between December 2006 and November 2009, 760 patients were randomly assigned, 612 in Italy (age ≤ 70 years) and 148 in Canada (without age limit).

All three trials were approved by the ethical committee of each participating institution. All patients signed written informed consent before enrollment in each trial. Treatment cycles were evaluable for our analysis if both toxicity evaluation and health-related quality-of-life (QoL) information were available in the study database. Analysis was limited to the first three treatment cycles.

Toxicity was prospectively collected by investigators. A paper case report form (CRF; sent by fax to coordinating center) was used for two trials (ELDA and GECO), and a Web-based CRF was available for the TORCH trial through the Web site of the coordinating center. The CRF (either paper or Web-based) was prepopulated with a specific list of adverse events (including those described in our analysis), and the worst grade (from 0 to highest one) for each adverse event was collected at the end of each cycle. Toxicity was coded according to the National Cancer Institute (NCI) Common Toxicity Criteria (version 2.0) in GECO and ELDA and the NCI Common Terminology Criteria for Adverse Events (CTCAE; version 3) in TORCH. Coding of the six toxicities considered in our analysis was the same in the two different versions of the NCI CTCAE used. For our analysis, any grade coded by the physician as > 0 was deemed “toxicity reported by the physician,” whatever the grade. For all six toxicities, according to the scales adopted, even the mildest toxicities should have been reported as grade 1.

At the end of each treatment cycle, patients completed the European Organisation for Research and Treatment of Cancer (EORTC) QoL questionnaires. The core questionnaire (QLQ-C30) was used in all the three trials, along with the lung cancer–specific module (QLQ-LC13) in the GECO and TORCH trials and the breast cancer–specific module (QLQ-BR23) in the ELDA trial. These instruments are designed to be completed by the patient. Anorexia, nausea, vomiting, constipation, and diarrhea are assessed by one item each in the QLQ-C30 questionnaire: items 13 (have you lacked appetite?), 14 (have you felt nauseated?), 15 (have you vomited?), 16 (have you been constipated?), and 17 (have you had diarrhea?). Hair loss is assessed by one item in QLQ-LC13 (item 39 [have you had hair loss?]) and one in QLQ-BR23 (item 34 [have you lost any hair?]). These questions specifically refer to the previous week. The items are scored in four categories (not at all, a little, quite a bit, or very much). All responses different from “not at all” (ie, a little, quite a bit, and very much) were pooled together as “any severity” or simply “any,” followed by the name of the adverse effect.

All results are reported per patient (ie, patient was unit of analysis), overall and separately by trial. A complementary per-cycle analysis, with cycle as unit of analysis, is reported in the Data Supplement. In the per-patient analysis, agreement between patient and physician evaluations was assessed by Cohen’s κ. Although there is no universal definition of the interpretation of κ values, according to Fleiss, κ values < 0.40 can be interpreted as poor agreement, values between 0.40 and 0.75 as moderate to good agreement, and values > 0.75 as excellent agreement. Under-reporting was calculated as the rate of cases where physicians reported grade 0 toxicity in all the cycles, of cases where patients reported toxicity in ≥ 1 cycle. In the per-cycle analysis, under-reporting was calculated as the rate of cycles where physicians reported grade 0, of cycles where patients reported toxicity. In the per-patient approach, two different analyses were performed, the first in all patients reporting any toxicity and the second limited to those patients reporting “very much” toxicity. A similar approach was used also in the per-cycle analysis.

Overall, of 1,459 patients enrolled onto the three randomized trials, 1,090 patients (enrolled by 78 institutions in two countries) were eligible for this analysis (because of availability of toxicity information and QoL questionnaire for ≥ one of first three cycles), for a total of 2,482 evaluable cycles. Compared with noneligible patients, those eligible more frequently had performance status of 0 and were of Canadian origin; also, eligible patients received a slightly higher mean number of cycles and, according to physician assessment, experienced more toxicity for all the analyzed items (Data Supplement). As outlined in Figure 1, 986 patients were assessed for cycle one, 840 for cycle two, and 656 for cycle three. The main characteristics of patients included are summarized in Table 1.

Patient reports of anorexia (of any severity) were documented in 679 (62.3%) of 1,090 patients, nausea in 654 (60.0%) of 1,089, vomiting in 283 (26.0%) of 1,090, constipation in 554 (51.0%) of 1,087, diarrhea in 388 (35.7%) of 1,088, and hair loss in 552 (50.8%) of 1,086. Physician reports of anorexia (of any grade) were documented in 202 (18.5%) of 1,090 patients, nausea in 488 (44.8%) of 1,089, vomiting in 256 (23.5%) of 1,090, constipation in 202 (18.6%) of 1,087, diarrhea in 248 (22.8%) of 1,088, and hair loss in 207 (19.1%) of 1,086.
Table 2 describes the agreement between patient reporting (any severity) and physician reporting (any grade) of toxicity in all patients. For the six toxicities, Cohen’s \(k\) ranged between 0.15 and 0.45, which can be interpreted as poor to moderate agreement. In all cases, percentages of toxicities reported by patients were higher than those reported by physicians. As shown in Figure 2, considering only patients who reported any toxicity, the proportion of under-reporting by physicians (ie, patients for whom physicians reported grade 0) was 74.4% for anorexia, 40.7% for nausea, 47.3% for vomiting, 69.3% for constipation, 50.8% for diarrhea, and 65.2% for hair loss. Some heterogeneity among trials was evident only for diarrhea, which was more frequently under-reported in GECO, and hair loss, which was less frequently under-reported in ELDA. Detailed numbers of the under-reporting, in the pooled data set and by trial and according to different treatment arms of each trial, are reported in the Data Supplement.

When examining only patients who reported “very much” toxicity in any cycle, the proportion of under-reporting by physicians was 50.0% for anorexia, 25.8% for nausea, 13.0% for vomiting, 44.2% for constipation, 24.1% for diarrhea, and 42.7% for hair loss. Detailed numbers are reported in the Data Supplement.

Analysis per cycle showed similar results, both for agreement and under-reporting. The estimates of under-reporting always exceeded those calculated in the per-patient analysis. Detailed numbers are reported in the Data Supplement.

**DISCUSSION**

Comparing the report of six symptomatic toxicities of anticancer treatment by patients and physicians in a large series of patients enrolled onto three randomized clinical trials, we found substantial rates of disagreement and under-reporting by physicians. As shown in Figure 2, considering only patients who reported any toxicity, the proportion of under-reporting by physicians (ie, patients for whom physicians reported grade 0) was 74.4% for anorexia, 40.7% for nausea, 47.3% for vomiting, 69.3% for constipation, 50.8% for diarrhea, and 65.2% for hair loss. Some heterogeneity among trials was evident only for diarrhea, which was more frequently under-reported in GECO, and hair loss, which was less frequently under-reported in ELDA. Detailed numbers of the under-reporting, in the pooled data set and by trial and according to different treatment arms of each trial, are reported in the Data Supplement.

When examining only patients who reported “very much” toxicity in any cycle, the proportion of under-reporting by physicians was 50.0% for anorexia, 25.8% for nausea, 13.0% for vomiting, 44.2% for constipation, 24.1% for diarrhea, and 42.7% for hair loss. Detailed numbers are reported in the Data Supplement.

Analysis per cycle showed similar results, both for agreement and under-reporting. The estimates of under-reporting always exceeded those calculated in the per-patient analysis. Detailed numbers are reported in the Data Supplement.
were explicitly asked to refer to the last week, whereas physicians were obviously requested to report all adverse effects experienced by the patient during the treatment cycle (lasting 4 weeks in ELDA trial and 3 weeks in GECO and TORCH trials). Many symptomatic toxicities—typically nausea and vomiting—are usually more frequent in the first days after treatment administration compared with the last week of the cycle. However, the possible bias arising from this inconsistency would be conservative, because the rate of symptoms reported by the patients should have been even higher if they reported on the whole duration of the treatment cycle rather than just the previous week, potentially further increasing the rate of physician under-reporting of toxicity.

The issue of under-reporting of toxicity in patients with cancer by physicians has been described in several studies. Basch et al prospectively compared reporting of symptom occurrence and severity by patients with cancer and clinicians (physicians or nurses) in an experimental study where clinicians were aware that their reports would be compared with patients’ reports. The agreement was quite low for symptoms prevalently caused by the disease and quite high, although not complete, for potential adverse effects. In another prospective blinded study, where clinicians were aware that their assessments would be compared with those of patients’, agreement was quite good for absence of toxicity but not satisfactory for grade of adverse effects. Recently, significant under-reporting of symptoms caused by androgen deprivation with or without docetaxel in patients with prostate cancer was also reported. Our data suggest that reporting by physicians of subjective toxicity in clinical trials may be not accurate enough. With >1,000 patients with cancer, evaluated for up to three cycles of treatment, our analysis was conducted in a large series of patients who were enrolled onto and prospectively observed in clinical trials. As specified in Patients and Methods, we limited our pooled analysis to three trials, selected because of uniform QoL instruments used and timing of administration. Other randomized trials coordinated by our clinical trials unit were not included in the analysis, essentially because of different QoL instruments or different timing of administration. However, our series is highly representative of the larger cancer population, including two common types of cancer, both elderly and younger patients, and both patients with early- and advanced-stage disease, with only slight differences in under-reporting by physicians in the three trials. This makes the generalizability of our findings quite high.

Several factors may explain under-reporting. First, on the physician side, there may be less attention paid to subjective toxicity, particularly for adverse effects that would not prompt treatment modification or supportive care, for those present before treatment initiation, and for those mild in severity. However, all six toxicities analyzed should have been coded at least as grade 1, even if mild in severity. Furthermore, to be conservative, we considered toxicity as correctly reported, whatever the grade specified by the physician. Second, at least in principle, physicians could have not reported an adverse event because they judged it unrelated to treatment. However, this could be true for some adverse events (like appetite loss) that are commonly related to cancer itself (at least in patients with advanced disease), but it is less likely for other adverse events (eg, nausea, vomiting, hair loss) that were probably treatment related in the majority of cases. Third, it is possible that clinicians could be less likely to report a toxicity that is largely expected with the drugs administered. However, at least in principle, the opposite could also be true, and clinicians
could be possibly more sensitive to largely expected toxicities. Looking at the under-reporting in the two arms of the TORCH trial (Data Supplement), the rate of under-reporting for nausea and vomiting is higher with erlotinib than with chemotherapy (latter being expected to produce more emesis), whereas the rate of under-reporting for diarrhea is higher with chemotherapy than with erlotinib (which includes diarrhea among its typical adverse events). Fourth, on the patient side, the attention paid in reporting verbally to the physician what was reported in single items of the QoL questionnaire might be suboptimal, with the potential assumption that the toxicity was communicated through the questionnaire, or because of some degree of self-censure, considering that communication with the physician also focuses on treatment benefit, which some patients may consider more relevant than some adverse effects. Fifth, the manner in which toxicity is explored during verbal communication between physicians and patients might be generic and not guided by the extensive list of potential adverse effects included in CRFs. Sixth, there might be under-reporting between the clinical files and the study CRF; we cannot quantify this phenomenon, because peripheral monitoring was not performed in these three trials.

Our analysis, based on a series of patients enrolled in 78 different centers, documents that on average, agreement is low when we consider physician report of subjective toxicities. We believe it is important to underline that under-reporting produces underestimation of the absolute rate of toxicity, which is highly relevant information for patients and their physicians in clinical practice, as well as regulatory bodies. Without an accurate estimation of the absolute rate of adverse events, the discussion of the benefit-risk ratio of treatment can be substantially biased, particularly for newer drugs entering standard clinical practice, with which the treating physician might have no or minimal previous experience.

Our findings emphasize the need for modifying the current system of toxicity assessment in clinical trials. Specifically, a collaborative reporting approach, where the patients directly report symptomatic toxicity information, which is then provided to clinicians to inform their CTCAE reporting, could improve the efficiency of reporting, and modern technologic supports (eg, tablets) could be used to facilitate patient reporting. The NCI has developed and tested an item bank and software system aimed at directly collecting information on symptoms and adverse events from patients with cancer participating in clinical trials. This system is called the patient-reported outcomes version of the CTCAE (PRO-CTCAE). Considering the substantial risk of under-reporting when the description of symptomatic toxicity is based only on physicians’ reports, we agree that the PRO-CTCAE system could represent a valid instrument to improve adverse event reporting data quality and comprehensiveness, promote communication between patients and clinicians, and improve clinical decision-making. Research is currently ongoing on the implementation of this tool in cancer clinical trials.

In conclusion, our analysis shows that even when prospectively collected within randomized controlled trials, subjective toxicities associated with anticancer treatments are at high risk of under-reporting by physicians. Our findings strongly support the incorporation of patient-reported information into toxicity reports in clinical trials.

REFERENCES


AUTHORS’ DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at www.jco.org.

AUTHOR CONTRIBUTIONS

Conception and design: Massimo Di Maio, Francesco Perrone
Provision of study materials or patients: Natasha B. Leighl, Francesco Nuzzo, Cesare Gridelli, Vittorio Gębba, Fortunato Cardiello, Sabino De Placido, Anna Ceribelli, Adolfo G. Favaretto, Andrea de Matteis, Ronald Feld, Charles Butts, Alessandro Morabito, Gaetano Rocco
Collection and assembly of data: Massimo Di Maio, Natasha B. Leighl, Francesco Nuzzo, Cesare Gridelli, Vittorio Gębba, Fortunato Cardiello, Sabino De Placido, Anna Ceribelli, Adolfo G. Favaretto, Andrea de Matteis, Ronald Feld, Charles Butts, Jane Bryce, Alessandro Morabito, Gaetano Rocco, Francesco Perrone
Data analysis and interpretation: Massimo Di Maio, Ciro Gallo, Natasha B. Leighl, Maria Carmela Piccirillo, Gennaro Daniele, Simona Signoriello, Alessandro Morabito, Francesco Perrone

Manuscript writing: All authors
Final approval of manuscript: All authors

© 2015 by American Society of Clinical Oncology.
Information downloaded from jco.ascopubs.org and provided by at Bibl.Centralizzata medicina e chirurgia on June 26, 2016. Copyright © 2015 American Society of Clinical Oncology. All rights reserved.

GLOSSARY TERMS

**health-related quality of life (HRQoL):** a broad multi-dimensional concept that usually includes self-reported measures of physical and mental health.

**patient-reported outcomes:** questionnaires used in a clinical setting to systemically collect information directly from the patient.

See Cancer.Net’s Extensive Video Library

Your patients can view videos of ASCO members discussing a variety of cancer topics at [www.cancer.net/videos](http://www.cancer.net/videos). Videos are also available through Cancer.Net’s YouTube channel at [www.youtube.com/cancerdotnet](http://www.youtube.com/cancerdotnet).
AUTHORS’ DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Symptomatic Toxicities Experienced During Anticancer Treatment: Agreement Between Patient and Physician Reporting in Three Randomized Trials

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO’s conflict of interest policy, please refer to www.asco.org/rwc or jco.ascopubs.org/site/ifc.

Massimo Di Maio
Consulting or Advisory Role: Eli Lilly, Boehringer Ingelheim

Ciro Gallo
No relationship to disclose

Natasha B. Leighl
No relationship to disclose

Maria Carmela Piccirillo
No relationship to disclose

Gennaro Daniele
No relationship to disclose

Francesco Nuzzo
No relationship to disclose

Cesare Gridelli
No relationship to disclose

Vittorio Gebbia
No relationship to disclose

Fortunato Ciardiello
No relationship to disclose

Sabino De Placido
No relationship to disclose

Anna Ceribelli
No relationship to disclose

Adolfo G. Favaretto
No relationship to disclose

Andrea de Matteis
No relationship to disclose

Ronald Feld
No relationship to disclose

Charles Butts
Honoraria: Pfizer, Merck, Boehringer Ingelheim, Bristol-Myers Squibb
Consulting or Advisory Role: Pfizer, Merck, Boehringer Ingelheim, Bristol-Myers Squibb

Jane Bryce
Consulting or Advisory Role: AstraZeneca

Simona Signoriello
No relationship to disclose

Alessandro Morabito
Consulting or Advisory Role: Pfizer

Gaetano Rocco
No relationship to disclose

Francesco Perrone
No relationship to disclose