Integrating Patient-Reported Outcomes Into Cancer Symptom Management Clinical Trials Supported by the National Cancer Institute—Sponsored Clinical Trials Networks

Jeff A. Sloan, Lawrence Berk, Joseph Roscoe, Michael J. Fisch, Edward G. Shaw, Gwen Wyatt, Gary R. Morrow, and Amylou C. Dueck

From the Department of Health Sciences Research, Mayo Clinic and Mayo Foundation, Rochester, MN; H. Lee Mofitt Cancer Center, University of South Florida, Tampa, FL; University of Rochester Cancer Center Research Base, Rochester, NY; The University of Texas M.D. Anderson Cancer Center, Houston, TX; Wake Forest University Community Clinical Oncology Program, Winston-Salem, NC; and College of Nursing, Michigan State University, East Lansing, MI.

Submitted May 23, 2007; accepted September 4, 2007.

Supported by numerous funding instruments of the National Cancer Institute including, but not limited to, all of the Community Oncology Research Base Infrastructure grants

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Address reprint requests to Jeff A. Sloan, PhD, Department of Health Sciences Research, Mayo Clinic, Harwick 8-29, 200 1st St SW, Rochester, MN 55905; e-mail: jsloan@mayo.edu.

© 2007 by American Society of Clinical Oncology

0732-183X/07/2532-5070/\$20.00 DOI: 10.1200/JCO.2007.12.7670

ABSTRACT

Patient-reported outcomes (PROs) are often the primary end point in symptom management trials. The scientific field of PROs is evolving, as evidenced by the US Food and Drug Administration's February 2007 release of a draft guidance for using PROs in effectiveness claims for drug labeling. This article presents issues encountered during use of PROs in National Cancer Institute–sponsored symptom management trials. Selected trials are presented that exemplify the challenges often seen in symptom management trials, and solutions are described. The examples presented include defining the appropriate end point, selecting and validating assessments, and answering the research questions through statistical analysis and interpretation. Progress has been made in addressing some of the unique challenges of PRO-based symptom management research. Many challenges still remain, but a foundational body of work now exists for more consistent and rigorous application of PROs into symptom management trials. There remains a need for more research in several methodologic aspects of design, analysis, and interpretation of symptom management trials.

J Clin Oncol 25:5070-5077. © 2007 by American Society of Clinical Oncology

INTRODUCTION

Patient-reported outcomes (PROs) and healthrelated quality-of-life (HRQOL) assessments are being applied more often in clinical research. 1,2 Recent events demonstrate the evolution of the field of PROs. The National Cancer Institute (NCI) –sponsored Cancer Outcomes Measurement Working Group published a book summarizing the history, challenges, and successes facing the assessment of PROs in cancer.3 Researchers for the North Central Cancer Treatment Group (NCCTG) published a 16chapter, two-part monograph concerning QOL assessment in cancer clinical trials.4 The US Food and Drug Administration (FDA) released a draft guidance for the use of PROs as end points to support product labeling.⁵ A special issue of the health economics journal Value in Health in December 2007 will contain a series of articles expanding on the recommendations in the FDA guidance document.6 Collectively, these works indicate that, although progress has been made, there remains a need for more research in several methodologic aspects of design, analysis, and interpretation of all trials that involve PROs.

This article will explore the use of PROs in cancer symptom management trials. The recent FDA guidance document defines PROs as "a measurement of any aspect of a patient's health status that comes directly from the patient (ie, without the interpretation of the patient's responses by a physician or anyone else). In clinical trials, a PRO instrument can be used to measure the impact of an intervention on one or more aspects of patients' health status. . ., ranging from the purely symptomatic (response of a headache) to more complex concepts (eg, ability to carry out activities of daily living), to extremely complex concepts such as QOL, which is widely understood to be a multidomain concept with physical, psychological, and social components."

Symptom control trials that involve end points like HRQOL and other PROs present unique challenges for design, 7-10 analysis, 11,12 and interpretation. 13,14 Selected examples of symptom management trials conducted by NCI's Community Clinical Oncology Program (CCOP) clinical trials network demonstrate strategies for managing some of these challenges. 15

SELECTED SYMPTOM MANAGEMENT TRIALS: DEFINING END POINTS—WHAT CAN BE MEASURED?

Trials for Bone and Brain Metastases

Symptom management trials present a unique challenge in that the precise end point of interest is often difficult to define. ^{3,5} Symptom management protocols, especially those with palliative intent, often are directed towards relieving or preventing patient suffering. Multiple trials developed by the Radiation Therapy Oncology Group (RTOG) for symptoms associated with bone metastases and brain metastases provide insight into the utility of PROs in these settings.

The first of six RTOG bone metastases treatment trials discussed here was conducted in 1974 (RTOG 7402). ¹⁶ This study randomly assigned patients to several different radiation therapy fractionation schedules, such as 30 Gy in 10 fractions or 40.5 Gy in 15 fractions. The primary end point of the trial was the patient-reported pain at the treatment site, defined as a score composed of the frequency of pain (on a 0 to 3 scale, representing less than once a day, once a day, several times a day, or constant, respectively) multiplied by the intensity of the pain (on a 0 to 3 scale, representing none, mild, moderate, or severe, respectively). Pain medication use was monitored on a similar scale. This scale was created for the trial, and although similar in form and in content to other assessment instruments validated in different patient populations, it underwent no formal validation before its use in the trial. This study was one of the first clinical trials to demonstrate the feasibility of collecting this information directly from cancer patients.

A similar patient-reported pain scale was then used in RTOG 7810, a phase I/II trial of hemibody irradiation for widespread bone metastases. ¹⁷ A new end point of net pain relief was introduced, which was defined as the duration of pain relief as normalized to the length of survival.

Pain relief treatments often consist of a mixture of interventions including narcotics, analgesics, and radiation therapy. A particular problem arises if pain is the only component of the primary end point because a patient experiencing no pain while taking high-dose narcotics will have a better pain score than a patient on no analgesics experiencing a low level of pain. RTOG also conducted a phase I/II study of fractionated hemibody irradiation (RTOG 8822). The primary end point was the maximum-tolerated dose of fractionated hemibody irradiation, as determined by the occurrence of any severe toxicity recorded by the NCI Common Toxicity Criteria (CTC). The only other outcome measures reported were time to new disease and time to new treatment. No PROs were included, despite all of the patients having the inclusion criteria of at least moderate pain on entry. The rationale for not documenting the response of the pain to treatment was not reported.

The next study for the palliation of bone metastases was RTOG 9714, a randomized trial of single- and multiple-fraction radiation therapy. A rigid end point of complete pain relief (ie, no pain and no narcotic use) was chosen as the PRO for this study. This trial abandoned the previous trials approach of focusing only on the treated sites and focused on a global end point, which may not be sensitive for a localized outcome. However, the limitations of the scales used in the trials previously described were avoided. The specific pain measure used was the Brief Pain Inventory. Secondary end points were also patient reported and included HRQOL, which was measured by the Functional Assessment of Cancer Therapy (FACT), and health status, which was measured by the Health Utilities Index. An ongoing

RTOG trial, RTOG 0517, is comparing the use of bisphosphonates with or without a single dose of a bone-absorbed radionuclide (strontium-89 or samarium-153). Again, an objective and composite end point was chosen for the primary end point (skeletal events, including re-treatment or fractures). However, PROs are included as secondary end points. The Brief Pain Inventory will measure pain, and the FACT-General and the EuroOoL will measure HROOL.²³

Brain metastases have also been a scientific focus within the RTOG. The first trials compared various fractionation schedules and had observer-oriented primary end points using the Neurologic Function Classification.²⁴ The Neurologic Function Classification is as follows: 1, able to work, neurologic findings minor or absent; 2, able to be at home, although nursing care may be required; neurologic findings present but not serious; 3, requires hospitalization and medical care, with major neurologic findings; and 4, requires hospitalization and is in a serious physical or neurologic state, including coma.

The rapid deterioration and short survival of most patients with brain metastases presents a challenge for the use of PROs in symptom management trials involving this patient population. Patient compliance decreases as their medical and neurologic condition deteriorates.

The RTOG has tried a variety of approaches to address this challenge. A phase II trial of twice-a-day radiation for patients with brain metastases, RTOG 8528, used an end point that was still observer oriented, using an improvement in the Neurologic Function Class, improvement of the brain imaging, or a decrease in needed corticosteroids. The follow-up phase III trial, RTOG 9104, used survival as its primary end point and included the Mini-Mental State Examination (MMSE) as a secondary end point. Although the MMSE was completed by the patient, it was scored by an observer and, therefore, is not, strictly speaking, a PRO.24 Subsequent RTOG research attempted to use end points that reflected patients' experiences. To achieve this goal, a battery of neurocognitive assessments previously validated in cancer patients was constructed to create instruments that were sufficiently sensitive and specific but not overly burdensome to the patient. RTOG 0018 was one phase II trial that evaluated the ability of patients with brain metastases to complete this neurocognitive testing battery. The results demonstrated that it was feasible for patients to complete the battery with a high degree of compliance.²⁵

The primary lesson learned from this series of trials is that outcome measures of palliative care protocols can benefit from obtaining data directly from patients. The challenge is to incorporate the patient's experience using measures that are valid and have a high degree of completion. Furthermore, the end points must sufficiently reflect patients' conditions so that they can adequately measure the effects of interventions. To date, the palliative care protocols have not completely achieved these goals but have made progress.

Describing the Relationship Between Overall HRQOL and Individual Symptoms: Megestrol Acetate for Anorexia in Head and Neck and Lung Cancer Patients Receiving Radiation Therapy

A challenge in defining end points for symptom management trials is the conceptual overlap between HRQOL and symptoms. A particular challenge arises in situations when studies alleviate symptoms but do not impact HRQOL, or vice versa.

The Comprehensive Cancer Center of Wake Forest University (CCCWFU) CCOP Research Base encountered this challenge in a series of their symptom management trials aimed at treating anorexia.

Two phase III randomized controlled trials were performed, one in lung cancer patients (CCCWFU 98199)²⁶ and the other in head and neck cancer patients (CCCWFU 97300),²⁷ to assess the effect of prophylactic megestrol acetate (MA) on weight concurrent with and for 12 weeks after definitive radiation therapy. The combined results from a meta-analysis of both trials regarding weight loss are presented in Figure 1. In contrast to a slight increase in mean weekly weight for patients receiving MA of 0.01 lb (95% CI, -0.39 to 0.4 lb; P = .98), patients receiving placebo experienced a mean weekly weight loss of 0.56 lb (95% CI, -0.98 to -0.15 lb; P = .001).

There were no significant differences in HRQOL between patients receiving placebo and those receiving MA. Furthermore, although overall HRQOL was similar between treatment groups, the symptom of anorexia was improved in the MA arm (P=.008). Although one symptom might be ameliorated, it may not be associated with an alteration in overall HRQOL. Study investigators concluded that prophylactic MA significantly reduces weight loss in lung and head and neck cancer patients receiving at least 50 Gy of radiation therapy.

Similar findings were reported by the NCCTG in trials of epoetin- α for alleviating fatigue in advanced cancer patients. ²⁸ Although hemoglobin values improved, PROs related to fatigue and HRQOL did not change. In several hot flash studies, the NCCTG has reported that patients who report reductions of greater than 50% in hot flash activity do not indicate a substantial change in HRQOL. ^{4,29}

Validating Diaries and Defining Clinical Significance: Vitamin E for Hot Flashes

Defining end points for symptom management trials often involves developing and validating a new measurement approach and defining what is meant by a clinically significant outcome. The NCCTG conducted a two-period, placebo-controlled, cross-over clinical trial involving 120 menopausal breast cancer survivors randomly assigned to receive either vitamin E or placebo for alleviation of hot flashes. This trial is representative of many trials that evaluate hot flash interventions using hot flash activity as reported by the patient (see Sloan et al²⁹ for a list). The trial was the first in a series and was initiated when no accepted assessment for hot flash activity existed. A diary was constructed and concurrently validated for the ability to consistently measure a woman's hot flash activity. 31,32

The study illustrates the need for considering clinical significance (beyond statistical significance) in reporting results from symptom

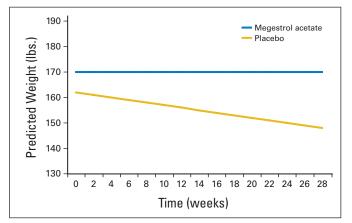


Fig 1. Predicted weight over time for megestrol acetate and placebo treatment groups in lung and head and neck cancer patients receiving radiation therapy.

management trials. The primary analysis for the first period of the cross-over design indicated that patients did not experience any difference in hot flash frequency between the placebo and vitamin E treatment arms (25% ν 22%, respectively; P = .90). Analysis of the full cross-over design, however, indicated that patients treated with vitamin E averaged one hot flash fewer per day than patients receiving placebo (roughly four ν five per day, respectively, from a baseline activity of about eight hot flashes per day; P < .05). The study was powered to detect a difference of at least one hot flash per day, which turned out to be close to the observed effect size. Although statistically significant, the results posed a question. Is it worth it to the individual to take the vitamin E and have one less hot flash? The investigators concluded that the "clinical magnitude of this reduction was marginal."³⁰ Although the P value is an important interpretation tool, the statistical analysis and results should reflect clinically meaningful outcomes and patient values.

SELECTING AND VALIDATING ASSESSMENTS: DOES THE MEASURE CAPTURE THE END POINT?

Selecting Appropriate Assessments: Ginger for Nausea

Once end points have been defined, another considerable challenge is to operationalize the definitions by selecting an appropriate method of assessment. The researcher must balance the idealized need to collect sufficient information against the reality that PRO assessment represents an additional burden to cancer patients. One way to deal with this challenge is exemplified by the University of Rochester Cancer Center (URCC) Research Base, which conducted a phase II/III randomized, controlled clinical trial of ginger for nausea caused by chemotherapy for cancer (URCC 0114). A 4-day home record of nausea was supplemented by the FACT-General QOL assessment and a symptom inventory to capture the adverse effects of ginger. Figure 2 displays the simple nausea scale used, which was easy for the patients to complete.

Figure 3 illustrates the straightforward manner in which the symptom inventory, adapted from one created at The University of Texas M.D. Anderson Cancer Center, was constructed.³³ Pain and fatigue are illustrated in Figure 3, but other symptoms were also captured in this manner. The only caveat to the inclusion of different symptoms in this inventory is that the terminology used for each symptom must be understandable by the patient. For example, one would not ask a patient to rate their peripheral neuropathy but, rather, the numbness or tingling in their hands and feet.

The URCC Research Base prefers assessment methods that keep measures and study design as short and as simple as possible. There is great utility in giving a symptom measure multiple times, and they have found that single items to assess symptom severity work best. Where there is not an obviously optimal symptom measure, the use of competing questionnaires assessing the symptom(s) of interest provide the required data to assess efficacy of the agent under study while also providing data to inform future trials of the same symptom. The URCC Research Base has published peer-reviewed empirical articles in the last 6 years that have included single-item assessments, typically at multiple time points. The URCC Research Base has provided ample evidence that simple symptom assessments allow for a direct approach to the evaluation of symptom ameliorative agents.

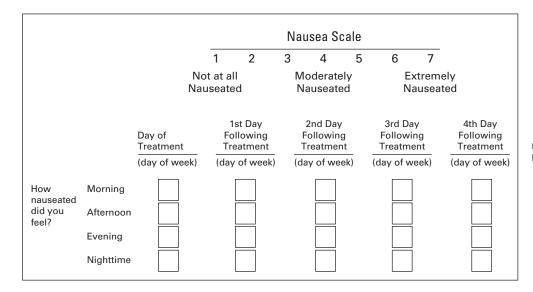


Fig 2. The nausea scale used by the University of Rochester Cancer Center Research Base.

Comparing Single-Item and Multiple-Item Assessments for the Same Symptom: Fatigue During Chemotherapy

Much has been written about the relative merits of single-item and multi-item assessments.³⁶ Single-item assessments represent an efficient method for assessing overall HRQOL domains but cannot capture the detail provided by multiple-item assessments. However, multiple-item assessments do not necessarily add clarifying information, as evidenced by a study involving the prevalence and treatment of fatigue in cancer patients receiving chemotherapy (URCC 1996P). 35 The primary objective in that study was to assess the degree to which an antidepressant drug (paroxetine) could reduce patient fatigue during chemotherapy treatment. Measures included the Multidimensional Assessment of Fatigue (MAF), the Profile of Mood States (POMS), the Fatigue Symptom Checklist, and the Center for Epidemiologic Studies Depression Scale (published elsewhere). The primary outcome measure reported was question 1 from the MAF, a 5-point semantic rating scale asking patients to indicate to what degree they had experienced fatigue during the previous week. Table 1 indicates that the correlation among study measures for fatigue was strong; most of the more complex measures are highly correlated with the simple, single MAF question. The single-item measure of fatigue also exhibited less overlap with the assessment of depression, as indicated by its lower correlations with the POM-Depression/Dejection and the Center for Epidemiologic Studies Depression Scale. The utility of a single-item fatigue assessment was confirmed in a separate phase III trial of modafinil using similar study measures.

Similar results have been seen for other end points in symptom management trials. 4,36 For example, normative data are now available for a single-item measure of overall HRQOL that has been used in numerous symptom management trials. 37 Research has further indicated that a score of 5 or less on the 0 to 10 scale used for this assessment of overall HRQOL is indication of the need for clinical intervention, that a movement of 2 points is a clinically meaningful change, and that a healthy individual typically will provide a score of 7 to 9. 12

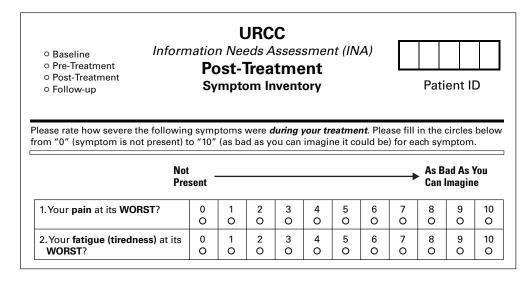


Fig 3. An example of a symptom inventory. The University of Rochester Cancer Center (URCC) Research Base Information Needs Assessment form.

Measure of Fatigue	Correlation				
	1	2	3	4	5
Multidimensional Assessment of Fatigue question 1	1				
Multidimensional Assessment of Fatigue total score	0.94	1			
3. Fatigue Symptom Checklist	0.67	0.73	1		
4. Profile of Mood States-Fatigue/Inertia	0.81	0.85	0.77	1	
5. Profile of Mood States-Depression/Dejection	0.47	0.52	0.61	0.60	1
Center for Epidemiologic Studies Depression Scale	0.60	0.67	0.71	0.70	0.76

Supplementing the CTC With PROs: Oxaliplatin-Related Peripheral Neuropathy

Symptom management trials have relied heavily on physicians to report patient symptoms, typically via the NCI CTC or, subsequently, the Common Terminology Criteria for Adverse Events. The CTC has a long history as a tool for reporting adverse events. The CTC specifies grading scales for hundreds of symptoms, typically assessed by a physician. For most end points relevant to a symptom management trial, the CTC provides definitions for an ordinal grading as follows: 0, none; 1, mild; 2, moderate; 3, severe; 4, life threatening; and 5, lethal.

A recent treatment study of the NCCTG dealing with oxaliplatinrelated peripheral neuropathy indicated that supplementing the CTC with simple PROs provided valuable clinically relevant information. The PROs involved numerical analog scale (NAS) assessments similar to the format already seen in Figure 3. The NAS asked patients to report their perceived level of peripheral neuropathy (numbness/tingling in their hands and feet) by circling an integer from 0 (not at all) to 10 (as bad as it can be). Six hundred ninety-six patients randomly assigned to receive oxaliplatin in the practice-changing NCCTG/Intergroup study N9741 provided data on the incidence and severity of peripheral neuropathy experienced via biweekly CTC assessments and NAS measurements taken every 12 weeks. ³⁸ Peripheral neuropathy via the NAS was defined as a worsening of 2 points from baseline. Time to onset of peripheral neuropathy via the NAS was defined as the time from baseline to the first 2-point worsening (from baseline) and was defined via the CTC as the time from baseline to the first report of grade 2 or higher (ie, grade 2+) peripheral neuropathy and the time from baseline to the first report of grade 3 or higher (ie, grade 3+) peripheral neuropathy. Two hundred seventy-six patients (40%) reported peripheral neuropathy via the NAS compared with only 99 patients (14%) with a reported incidence of CTC grade 3+ peripheral neuropathy. Figure 4 further demonstrates that the time to onset of peripheral neuropathy via the NAS was substantially less than the time to onset of CTC grade 3+ peripheral neuropathy. Hence, the NAS was able to record more often and more quickly when a patient should be considered for oxaliplatin dose modification.

A grade 3+ CTC report is typically the predicating factor in considering oxaliplatin dose modification. If one were to use incidence of a grade 2+ CTC peripheral neuropathy as an indicator event, this would be giving the CTC an unfair advantage because a grade 2 event is not typically considered as a trigger for dose modification. Even with this modification to the CTC usage, the time to onset of CTC grade 2+ peripheral neuropathy was still substantially longer than time to onset of peripheral neuropathy via the NAS for patients

with onset within the first 6 months of random assignment (Fig 4). The incidence rate of recording a CTC grade 2+ peripheral neuropathy was 37% and, therefore, closer to the incidence rate of 40% observed by the NAS. However, there was, at best, moderate concordance between the CTC and NAS approaches to identifying which patients were actually experiencing troublesome peripheral neuropathy (65%, Cohen's $\kappa = .26$). Hence, the extra people who were identified by the report of a CTC grade 2 event were, for the most part, patients who had not perceived a clinically meaningful change in their peripheral neuropathy via the NAS. The CTC and NAS PRO would seem to be measuring two different aspects of oxaliplatin-related peripheral neuropathy, with the NAS being more sensitive to the patient's perceptions. This is clinically important for potentially identifying a need for dose adjustment earlier or considering a prophylactic intervention to prevent the peripheral neuropathy from escalating to the level of a serious adverse event.

SELECTED SUCCESSFUL SYMPTOM MANAGEMENT TRIALS: ANALYZING AND INTERPRETING RESULTS

Listening to Patient Comments: Sucralfate for Stomatitis

Having defined appropriate end points and chosen the optimal assessment approach, the final major challenge remains analyzing and interpreting the results. Standard statistical methods deal with many

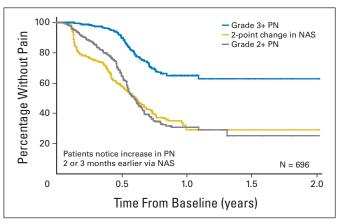


Fig 4. Time to peripheral neuropathy (PN): National Cancer Institute Common Toxicity Criteria (CTC) versus the numerical analog scale (NAS).

of the research questions posed by symptom management trials.^{3,9,12} Nonetheless, several statistical considerations involved in symptom management trials require particular attention. The intent-to-treat principle of including all patients in the analysis who are subjected to treatment and yet do not provide data is a particularly challenging aspect of symptom management trials. 9,11,12 This is exemplified in a trial reported by Martenson et al³⁹ of a placebo-controlled study of 100 patients testing sucralfate for alleviating stomatitis. In the initial analysis, P = .06 in favor of sucralfate when comparing the maximum grade of stomatitis for patients receiving sucralfate versus placebo. However, twice as many patients went off study early on the sucralfate arm. Investigation of the patient comments among those who did not complete the trial revealed that all but three patients on the sucralfate arm went off study because the treatment made them gag. If these patients were added to the analysis as treatment failures, the primary analysis P value would still be .06, but now, it would favor placebo over sucralfate. This study provided the lesson that, in symptom management trials, it is important to make sure you are asking the appropriate question.

Incorporating Mood As a Covariate: Antidepressants for Hot Flash Activity

Assessing the burden of symptoms in cancer patients without considering depression is difficult. Early in the 1990s, it became apparent that newer antidepressants that selectively inhibit serotonin reuptake (selective serotonin reuptake inhibitors) might reduce hot flash activity. Loprinzi et al⁴⁰ reported that venlafaxine administered in doses less than that approved to treat depression reduced hot flashes. The issue was differentiating activity specific to hot flash reduction from perception or mood. An important covariate, the Beck Depression Inventory, was used to address this issue. The adjusted analysis indicated that the mood changes were both statistically and clinically nonsignificant. Therefore, the effect of the venlafaxine on hot flash activity could be reasonably ascribed to true impact on the number of hot flashes experienced.

Consistency of PRO assessments over the series of hot flash studies carried out by the NCCTG allowed for a meta-analysis that yielded several implications for future study design.²⁹ Specifically, a scientifically supportable response criterion based on empirical data was defined as a 50% reduction in hot flash activity for individual patients. Furthermore, a sample size of only 25 patients was determined to be efficient for pilot hot flash studies.

Examining Symptom Clusters: Donepezil for Cognitive Function in Irradiated Brain Tumor Patients

Recently, researchers have attempted to analyze clusters of end points in an effort to interpret the multiplicative nature of PROs. Cognitive function, for example, is a simple collective concept to express but is a complex composite end point when incorporated into a clinical trial.

The Wake Forest University CCOP Research Base encountered this challenge when they conducted an open-label phase II study, CCCWFU 97100, which treated irradiated brain tumor patients with donepezil, an FDA-approved reversible acetylcholinesterase inhibitor used to treat mild to moderate dementia of the Alzheimer's type. ⁴¹ In addition to extensive cognitive assessments, this trial included the FACT-Brain for assessing HRQOL ⁴² and the POMS for assessing mood, ⁴³ illustrating the need for multiple symptom assessments to answer a composite research hypothesis.

Overall HRQOL as measured by the FACT-Brain significantly improved (ie, increased) from baseline to 24 weeks (mean score, 124 at baseline ν 134 at 24 weeks). Total score on the POMS also significantly improved (ie, decreased) from baseline to 24 weeks, indicating an improved mood (mean score, 47 at baseline ν 30 at 24 weeks). The various cognitive function assessments also significantly improved between baseline and week 24. There were no significant changes in Karnofsky performance status, global cognitive function (as measured by the MMSE), or executive function (as measured by the Trail-Making Test Part B). The most common adverse events were fatigue, insomnia, and diarrhea.

Results from the donepezil study led to methodologic research in a patient population of symptom clusters, defined as the coexistence of two or more symptoms as a result of cancer and its treatment.⁴⁴ The donepezil study database allowed researchers to characterize the symptom clusters that occur in irradiated brain tumor survivors.⁴⁵ Factor analysis, multidimensional scaling, and multivariate cluster demonstrated that three distinct symptom clusters-mood, cognition, and energy-exist in irradiated brain tumor survivors. The mood cluster included depression, anger, anxiety, and confusion. The cognitive function cluster included inability to concentrate, read, remember, and find the right words. The energy cluster included lack of energy, fatigue, and frustration. Further symptom treatment studies are planned in irradiated brain tumor patients to focus on therapeutic interventions for an entire symptom cluster rather than the traditional emphasis on a single symptom. These trials, also from Wake Forest, may include the use of multiple interventions (pharmacologic and behavioral) rather than a single pharmacologic intervention.

Dealing With Missing Data: Radiofrequency Ablation for Bone Metastases

An important consideration for conducting a successful symptom management trial is the proper management of missing data. In recent years, several methods have been advanced to both prevent and account for missing data in an experimental design. 46,47 Despite these statistical advances, missing data continue to be a major problem facing symptom management trials. PROs, which are reliant on patients being able and willing to provide a response, are often the primary end point in symptom management trials. The best method for handling missing data is prevention, 46 which requires careful consideration during the design and implementation stages of a clinical trial. Good planning is essential to avoid missing data, paying particular attention to the data collection burden of instrumentation selected. Research has shown that the amount of missing data is directly proportional to the number of items the patients is asked to complete. 47

Once missing data are present, the key issue is whether a systematic bias in the pattern of the missing data is associated with a particular treatment or whether the pattern of missing data is consistent across treatments. For example, experience has demonstrated that the best performers are the patients who tend to provide more complete data in symptom management trials. Patients who do not provide complete data often do so because they become too ill to participate. As a result, the PRO data in symptom management trials tend to overestimate the well-being of the patient population. Hence, analysis of PRO data in symptom management trials must take the presence of missing data into consideration. Various imputation methods have been proposed to deal with this issue.^{46,47}

A symptom management trial that dealt with missing data issues examined radiofrequency ablation for pain from bone metastases.⁴⁸ Missing data were anticipated as a result of the nature of the intervention and the desire to track pain longitudinally. A number of preventive measures were put in place. First, the number of assessments involved was scrutinized, and only items that were absolutely essential were included. Second, study assistants specified at each clinical site collected the PRO assessment data. Third, the assessments were organized into carefully designed high-quality print booklets that included a message to the participants from the investigator and a contact number for them to use for any questions or concerns. These efforts kept the missing data well below 10% for any end point. Current research suggests that keeping missing PRO data below the 10% threshold is essential for confidence in the findings of most studies, assuming that appropriate statistical imputation techniques are applied.⁴⁹ This conclusion reinforces the importance of careful planning and execution of PROs in clinical trials to minimize missing data in the first place.

CONCLUSION

The goal of this article was to examine the application of PROs in symptom management cancer clinical trials. Examples were drawn from specific NCI-sponsored symptom management trials that either involved methodologic techniques or furthered the knowledge base of

REFERENCES

- 1. Bottomley A, Vanvoorden V, Fletchner H, et al: The challenges and achievements involved in implementing QOL research in cancer clinical trials. Eur J Cancer 39:275-285, 2003
- 2. Osoba DA: Taxonomy of the uses of health-related quality-of-life instruments in cancer care and the clinical meaningfulness of the results. Med Care 40:III31-III38, 2002
- 3. Lipscomb J, Gotay CC, Snyder C: Outcomes Assessment in Cancer: Measures, Methods, and Applications. Cambridge, United Kingdom, Cambridge University Press, 2005
- **4.** Sloan JA: Challenges in QOL assessment: Overview and introduction. Curr Probl Cancer 29: 274-277, 2005
- **5.** US Food and Drug Administration: Patient-reported outcome measures: Use in medical product development to support labeling claims. http://www.fda.gov/cder/guidance/5460dft.pdf
- 6. Sloan JA, Halyard M, Frost M, et al: The FDA guidance for patient-reported outcomes: What does it mean and how will it impact clinical trials? Val Health 10:S59-S146, 2007
- **7.** Fayers PM, Hopwood P, Harvey A, et al: Quality of life assessment in clinical trials: Guidelines and a checklist for protocol writers—The U.K. Medical Research Council Experience. Eur J Cancer 33:20-28, 1997
- **8.** Gotay CC, Korn EL, McCabe MS, et al: Quality-of-life assessment in cancer treatment protocols: Research issues in protocol development. J Natl Cancer Inst 84:575-579, 1992
- 9. Fairclough DL: Design and Analysis of Quality of Life Studies in Clinical Trials. New York, NY, CRC Press, 2002

applications of PROs for cancer-related symptoms. These areas of research continue to evolve.

The clinical trials described in this article demonstrate that incorporating PROs in symptom management trials presents challenges in study design, analysis, and interpretation. Many challenges still remain, but a foundational body of work now exists for more consistent and rigorous application of PROs in future symptom management trials. Recent publications and workshops, such as Patient-Reported Outcomes Assessment in Cancer Trials, have attempted to delineate the key issues now facing researchers who want to incorporate the patient's subjective experience into cancer clinical trials.^{3,5} Further methodologic advances similar to those presented in this article are needed to ensure that PROs have the requisite scientific integrity for continued and expanded use in symptom management clinical trials.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Manuscript writing: Jeff A. Sloan, Lawrence Berk, Joseph Roscoe, Michael J. Fisch, Edward G. Shaw, Gwen Wyatt, Gary R. Morrow, Amylou C. Dueck

- 10. Dancey J, Zee B, Osoba D, et al: Quality of life scores: An independent prognostic variable in a general population of cancer patients receiving chemotherapy—The National Cancer Institute of Canada Clinical Trials Group. Qual Life Res 6:151-158, 1997
- 11. Fayers P, Hays R: Assessing Quality of Life in Clinical Trials. New York, NY, Oxford University Press 2005
- 12. Sloan JA, Dueck A: Issues for statisticians in conducting analyses and translating results for quality of life end points in clinical trials. J Biopharm Stat 14:73-96, 2004
- 13. Staquet MJ, Berzon RA, Osoba D, et al: Guidelines for reporting results of quality of life assessments in clinical trials, in Staquet MJ, Hays RD, Fayers PM (eds): Quality of Life Assessment in Clinical Trials. New York, NY, Oxford University Press, 1998
- **14.** Symonds T, Berzon R, Marquis P, et al: Clinical significance of quality of life results: Practical considerations for specific audiences. Mayo Clin Proc 77:572-583, 2002
- 15. Buchanan DR, White JD, O'Mara AM, et al: Research-design issues in cancer-symptom-management trials using complementary and alternative medicine: Lessons from the National Cancer Institute Community Clinical Oncology Program experience. J Clin Oncol 23:6682-6689, 2005
- **16.** Tong D, Gillick L, Hendrickson FR: The palliation of symptomatic osseous metastases: Final results of the study by the Radiation Therapy Oncology Group. Cancer 50:893-899, 1982
- 17. Salazar OM, Rubin P, Hendrickson FR, et al: Single-dose half-body irradiation for palliation of multiple bone metastases from solid tumors: Final Radiation Therapy Oncology Group report. Cancer 58: 29-36, 1986

- **18.** Scarantino CW, Caplan R, Rotman M, et al: A phase I/II study to evaluate the effect of fractionated hemibody irradiation in the treatment of osseous metastases: RTOG 88-22. Int J Radiat Oncol Biol Phys 36:37-48, 1996
- **19.** Hartsell WF, Scott CB, Bruner DW, et al: Randomized trial of short- versus long-course radiotherapy for palliation of painful bone metastases. J Natl Cancer Inst 97:798-804, 2005
- **20.** Cleeland CS, Ryan KM: Pain assessment: Global use of the Brief Pain Inventory. Ann Acad Med Singapore 23:129-138, 1994
- **21.** Cella DF, Tulsky DS, Gray G, et al: The Functional Assessment of Cancer Therapy scale: Development and validation of the general measure. J Clin Oncol 11:570-579, 1993
- 22. Feeny D, Furlong W, Barr RD: Multiattribute approach to the assessment of health-related quality of life: Health Utilities Index. Med Pediatr Oncol 1:54-59, 1998 (suppl)
- 23. Krabbe PF, Peerenboom L, Langenhoff BS, et al: Responsiveness of the generic EQ-5D summary measure compared to the disease-specific EORTC QLQ C-30. Qual Life Res 13:1247-1253, 2004
- **24.** Borgelt B, Gelber R, Kramer S, et al: The palliation of brain metastases: Final results of the first two studies by the Radiation Therapy Oncology Group. Int J Radiat Oncol Biol Phys 6:1-9, 1980
- **25.** Regine WF, Schmitt FA, Scott CB, et al: Feasibility of neurocognitive outcome evaluations in patients with brain metastases in a multi-institutional cooperative group setting: Results of Radiation Therapy Oncology Group trial BR-0018. Int J Radiat Oncol Biol Phys 58:1346-1352, 2004
- **26.** McQuellon RP, Moose DB, Russell GB, et al: Supportive use of megestrol acetate (Megace) with head/ neck and lung cancer patients receiving radiation therapy. Int J Radiat Oncol Biol Phys 52:1180-1185, 2002

JOURNAL OF CLINICAL ONCOLOGY

- 27. Farmer M, Case D, Lesser G, et al: A phase III double blind placebo controlled prospective randomized trial on the effect of megestrol acetate on weight and health-related quality of life in lung cancer and head/neck cancer patients receiving definitive radiation therapy. Int J Radiat Oncol Biol Phys 63:S77, 2005 (abstr)
- **28.** Witzig TE, Silberstein PT, Loprinzi CL, et al: Phase III, randomized, double-blind study of epoetin alfa compared with placebo in anemic patients receiving chemotherapy. J Clin Oncol 23:2606-2617, 2005
- 29. Sloan JA, Loprinzi CL, Novotny PJ, et al: Methodological lessons learned from hot flash studies. J Clin Oncol 19:4280-4290. 2001
- **30.** Barton D, Loprinzi C, Quella S, et al: Prospective evaluation of vitamin E for hot flashes in breast cancer survivors. J Clin Oncol 16:495-500, 1998
- **31.** Finck G, Barton DL, Loprinzi CL, et al: Definitions of hot flashes in breast cancer survivors. J Pain Symptom Manage 16:327-333, 1998
- **32.** Carpenter JS: State of the science: Hot flashes and cancer. Part 1: Definition, scope, impact, physiology, and measurement. Oncol Nurs Forum 32:959-968, 2005
- **33.** Cleeland CS, Mendoza TR, Wang XS, et al: Assessing symptom distress in cancer patients: The M.D. Anderson Symptom Inventory. Cancer 89: 1634-1646. 2000
- **34.** Morrow GR, Roscoe JA, Hickok JT, et al: Nausea and emesis: Evidence for a biobehavioral perspective. Support Care Cancer 10:96-105, 2002

- **35.** Morrow GR, Hickock JT, Roscoe JA, et al: Differential effects of paroxetine on fatigue and depression: A randomized, double-blind trial from the University of Rochester Cancer Center Community Clinical Oncology Program. J Clin Oncol 21: 4635-4641, 2003
- **36.** Sloan JA, Aaronson N, Cappelleri JC, et al: Assessing the clinical significance of single items relative to summated scores. Mayo Clin Proc 77: 479-487, 2002
- **37.** Locke D, Sloan JA, Novotny PJ, et al: Validation of and normative data for simple single-item numerical analogue self-assessments for quality of life. J Pain Symptom Manage [epub ahead of print on August 20, 2007]
- **38.** Goldberg RM, Sargent DJ, Morton RF, et al: A randomized controlled trial of fluorouracil plus leucovorin, irinotecan, and oxaliplatin combinations in patients with previously untreated metastatic colorectal cancer. J Clin Oncol 22:23-30, 2004
- **39.** Martenson JA, Bollinger JW, Sloan JA, et al: Sucralfate in the prevention of treatment-induced diarrhea in patients receiving pelvic radiation therapy: A North Central Cancer Treatment Group phase III double-blind placebo-controlled trial. J Clin Oncol 18:1239-1245, 2000
- **40.** Loprinzi C, Kugler J, Sloan J, et al: Venlafaxine in management of hot flashes in survivors of breast cancer: A randomised controlled trial. Lancet 356: 2059-2063, 2000
- **41.** Shaw EG, Rosdhal R, D'Agostino RB Jr, et al: A phase II study of donepezil in irradiated brain

- tumor patients: Effect on cognitive function, mood, and quality of life. J Clin Oncol 24:1415-1420, 2006
- **42.** Weitzner MA, Meyers CA, Gelke C, et al: The Functional Assessment of Cancer Therapy (FACT) scale: Development of a brain subscale and revalidation of the general version (FACT-G) in patients with primary brain tumors. Cancer 75:1151-1161, 1995
- **43.** Curran SL, Andrykowski MA, Studts JL: Short Form of the Profile of Mood States (POMS-SF): Psychometric information. Psychol Assess 7:80-83, 1995
- **44.** Dodd MJ, Miaskowski C, Lee KA: Occurrence of symptom clusters. J Natl Cancer Inst Monogr 32:76-78, 2004
- **45.** Saconn PA, Ip E, Rapp S, et al: Symptom clusters in irradiated brain tumor survivors. J Clin Oncol 24:488s, 2006 (suppl; abstr 8581)
- **46.** Fairclough D, Peterson HF, Chang V: Why are missing quality of life data a problem in clinical trials of cancer therapy? Stat Med 17:667-677, 1998
- **47.** Huntington JL, Dueck A: Handling missing data. Curr Probl Cancer 29:317-325, 2005
- **48.** Goetz MP, Callstrom MR, Charboneau JW, et al: Percutaneous image-guided radiofrequency ablation of painful metastases involving bone: A multicenter study. J Clin Oncol 22:300-306, 2004
- **49.** Dillman DA: Mail and Internet Surveys: The Tailored Design Method. New York, NY, John Wiley & Sons, 2000
