Outcomes Research in Cancer Symptom Management Trials: The Radiation Therapy Oncology Group (RTOG) Conceptual Model

Deborah Watkins Bruner

The Radiation Therapy Oncology Group (RTOG) Health Services Research and Outcomes (HSRO) Committee aims to guide the study of the interactions among clinical, humanistic, and economic variables that optimize patient outcomes on clinical trials. To guide this work, the RTOG Outcomes Model was developed. Within this framework, measurement focuses primarily on patient-reported outcomes (PROs). In the examples presented, these outcomes have served to better quantify the benefit of one therapy over alternative therapies, as in the example of multimodality therapy for lung cancer, and to add evidence to clinical outcomes when clinical outcomes alone have not been strong enough to change clinical practice, as in the example of palliative radiotherapy for painful bone metastasis. The unique contribution to the RTOG of the HSRO Committee is the selection and use of PRO measures that give "voice" to the patient in clinical trials as well as provide data to better manage symptoms.


The science of cancer symptom management has come far in a relatively short time. Twenty-five years ago, the literature on symptom management was still heavily focused on palliative care for the terminally ill, and the measure of symptom management success was still commonly assessed by clinician evaluation of toxicities (1). In the 1980s, investigators began advocating for measures that assess the patients' perspective of symptoms and psychological distress (2,3). These measures began to develop into patient-reported health-related quality of life (QOL) measures, first with simple visual analogue scales (4,5) and interview questionnaires (6) to more sophisticated instrument development (7–9).

In the National Cancer Institute (NCI) Clinical Trials Cooperative Groups (CTCG), measuring the impact symptoms had on QOL was actually spurred by the Food and Drug Administration (FDA). The FDA stimulated interest in QOL outcomes in 1985, when it published approval requirements for anticancer drugs that included favorable effects on survival and/or QOL (10). The Division of Cancer Treatment of the NCI followed the FDA's lead by revising its mission statement from 1988 to declare: "Research aimed at improving survival and QOL for persons with cancer is of the highest priority to The Cancer Therapy Evaluation Program [CTEP]" (11). The FDA has again taken the lead in suggesting regulations for the methodology of QOL and patient-reported outcomes (PROs) research that will have far-reaching consequences in this field with its Draft Guidance on Patient Reported Outcomes (12).

Following their 1988 mission statement, CTEP actively encouraged the national CTCG to incorporate QOL endpoints into their clinical trials. The Radiation Therapy Oncology Group (RTOG) readily complied, and the RTOG QOL Committee was established in 1989. The purpose of this paper is to review a conceptual framework for the incorporation of outcomes research in cancer symptom management trials developed by the RTOG. In addition, we will briefly review governmental edicts and the state of the art of outcomes research that have had an impact on incorporating PROs into symptom management trials in the NCI-sponsored CTCG.

Methods

The RTOG QOL Committee was formed as a complement to the primary Site and Modality Committees. The mission of the QOL Committee was to assess health-related QOL outcomes related to radiation therapy, multimodality therapy, and symptoms.

However, as RTOG's work in QOL and symptom management grew, so did the outcomes of interest. In addition to QOL, other PROs (for example, behavioral, cognitive, and preferences) began to be studied. These PROs expanded the more traditional measurement of outcomes evaluated in RTOG such as mortality, symptoms measured with traditional physical and biomedical examinations, and toxicity as measured with physician-rated toxicity-grading systems.

In addition to the evaluation of PROs related to various treatments, interventional studies to manage or prevent symptoms have always been a major focus of the RTOG Community Clinical Oncology Program and continues with current investigations of agents for the treatments of diarrhea, erectile dysfunction, mucositis, and pulmonary fibrosis. But it was the additional interest in patient preferences, utilities, quality-adjusted life years, and economic outcomes that spurred RTOG to develop a new paradigm for assessing outcomes related to both symptom and treatment trials.

In 2000, the RTOG Outcomes Committee (now the Health Services Research and Outcomes [HSRO] Committee) was formed as a complement to the primary Site and Modality Committees. Its mission is to further the work of the RTOG by studying endpoints other than, or in conjunction with, mortality, morbidity, and QOL. The HSRO Committee initially defined its vision as

Correspondence to: Deborah Watkins Bruner, RN, PhD, School of Nursing, University of Pennsylvania, Nursing Education Building, Rm 352, 420 Guardian Dr, Philadelphia, PA 19104-6096 (e-mail: wbbruner@nursing.upenn.edu).

DOI: 10.1093/jncimonographs/lgm004
© The Author 2007. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oxfordjournals.org.
the discovery of the interactions among clinical, humanistic, and economic variables that optimize the outcomes and use of resources for defined populations (13). The development and implementation of the RTOG Outcomes Model that guides the assessment and interaction of these three variables (identified as a triad in the Model, Fig. 1) are described in detail elsewhere (13,14).

Briefly, thinking of the model as a triangle with the clinical endpoints at the apex and the humanistic endpoints at the base point on the right and the economic endpoints at the base point on the left, use of the model guides the measurement and combination of all three endpoints which can be analyzed separately and in combination (as in a quality-adjusted life year and a cost–utility equation). Within this framework, measurement focuses primarily but not exclusively on PROs.

**Results**

RTOG’s first study of QOL outcomes was RTOG 9020, a phase II trial of external beam radiation therapy plus etanidazole for locally advanced prostate cancer. The study was designed to assess institutional and patient compliance with patients’ self-reported QOL questionnaires (e.g., the Functional Assessment Cancer Therapy [FACT]; The Sexual Adjustment Questionnaire) in RTOG clinical trials (15). The authors concluded that PRO assessments were attainable in RTOG clinical trials and that compliance rates were acceptable at the initial time point and at 3-month follow-up. A secondary aim of the study was to compare patients’ self-report of symptoms, as measured by specific items on a QOL instrument, with medical professional ratings of the same symptoms using the RTOG acute toxicity rating scale. Disagreement between patients’ self-report of symptoms on the QOL scale and medical professional ratings on the RTOG acute toxicity rating scales of the same symptoms (e.g., sexual function) ranged from 13% to 45% at 3 months.

Congruent with the first RTOG QOL study, other RTOG studies have documented differences in clinician- and patient-reported outcomes of the same symptoms. Bruner et al. (16) examined sexual outcomes in 471 prostate cancer patients and found that physician and patient assessments of the patient’s ability to have an erection differed up to 47% of the time. More recent RTOG studies have continued to report discrepancies between physician- and patient-reported outcomes. A phase III study of Amifostine (RTOG 9801) for mucosal protection for patients with metastases from breast or prostate cancers. The study had as its primary objective to demonstrate that the combination therapy as compared with monotherapy (14). However, at the humanistic point of the triad, two quality-adjusted survival analyses noted that for specific subgroups, particularly patients more than 70 years of age compared with younger patients, the same level of benefit of combined therapies was not demonstrated due to increased toxicities (18,19). This raised the hypothesis for a symptom management trial of esophagitis (17). To assess all three points of the triad, the clinical and humanistic outcomes (quality-adjusted survival) were combined with economic outcomes in a cost–utility analysis that indicated that induction chemotherapy followed by radiation therapy was the most cost-effective therapy, as compared with several alternatives (20).

A more recent example of a symptom control trial prospectively guided by the RTOG Outcomes Model was RTOG 9714, a phase III trial to investigate whether 8 Gy delivered in a single treatment fraction provides pain and narcotic relief (the clinical and humanistic outcomes) that are equivalent to that of the standard treatment course of 30 Gy delivered in 10 treatment fractions for a period of 2 weeks, for patients with painful bone metastases from breast or prostate cancers. The study had as its secondary aim the assessment of differences between arms in health-related QOL as measured by patient-reported measures of domain-specific (i.e., pain) and more global QOL (i.e., FACT) and health state valuations (i.e., utilities). (Utilities are a measure of how a particular therapy or symptom is valued on a scale ranging from 0, indicating death or worse possible health state, to 1, indicating perfect health.) The primary outcome was documented as a complete or partial improvement in pain in 66% of patients on both arms. Pain and narcotic relief were equivalent for both 30 Gy in 10 fractions and 8 Gy in 1 fraction (21). The utilities supported the primary outcome, indicating that there was no difference between the two treatment arms at baseline or at 3 months in health-related QOL valuations (22). The utility scores will be combined with survival in a quality-adjusted survival analysis, and these results are further planned for use in a cost–utility analysis. The quality-adjusted survival analysis and cost–utility analysis are needed to further support the clinical outcomes because similar past clinical findings were discounted by many physicians who continued to use 30 Gy based on unsubstantiated claims that the longer course of radiotherapy may...

---

**Fig. 1.** RTOG Outcomes Model—research endpoints. PROs = patient-reported outcomes; QOL = quality of life; QAS = quality-adjusted survival. Adapted and reprinted from Watkins-Bruner D, Berk L, Bondy M, Kachnic L, Konski A, Layne E, et al.; Outcomes Committee: RTOG Core Grant. Int J Radiat Oncol Biol Phys 2001; 51:66–74. Copyright 2001, with permission from Elsevier.
improve both symptom-specific and global QOL and therefore be cost-effective in the long run.

**Discussion**

Findings from RTOG 9020 and 9801 suggest that patient-reported assessments of symptoms appear to pick up some results, particularly subjective symptoms (i.e., pain, sexual function, swallowing, etc.), better than the physician-reported toxicity ratings. PROs are important value-added measures that complement toxicity-grading scales. Yet, at least in most of the CTCG, there is still some debate on their use, and when issues of cooperative group resources or patient burden are of concern, PROs may often be the first measures to be deleted from a study. PRO instruments have been developed to assess symptoms and other outcomes from the patient perspective, which may or may not be related to the toxicities. The patients' perspective should be paramount when assessing any potentially subjective outcomes of treatments or symptom interventions.

Whether to measure PROs or physician-rated toxicities should never be an either/or question, just as it is not an either/or question to measure mortality and morbidity in a trial where these outcomes are of interest. The more pertinent question is which PRO measure(s) would provide the best information for improving care in the population under study. The now numerous PROs to choose from leave us to consider questions regarding which provide the best data for future interventions (such as symptom-specific measures or more global measures of health status or QOL), which have been validated in the population under study, or in studies where validated measures are unavailable, what reasonable options (versus doing nothing) are available for measurement.

Further, continued arguments or innuendo of the interpretability or meaning of PROs is unfounded. Well-designed PRO instruments undergo rigorous psychometric testing that most of the cancer-related toxicity scales, including the CTC, have not done. In fact, in comparison, some common biomedical measures have frequently demonstrated lesser degrees of reliability than most commonly used PRO measures within the CTCG. For example, in studies to determine the accuracy and reliability of blood pressure readings with sphygmomanometers used in emergency medical services, it was found that if accuracy and reliability were considered together, a total of 73% (109/150) of the devices failed one or both of the criteria and that the reproducibility of vital sign measurements may be further limited by significant interobserver variability (23,24). One study of interrater reliability of chest radiographic results collected in a multisite clinical trial showed moderate agreement for most findings, with the best correlation reported for the presence of bronchovascular markings and/or reticular densities addressed as a composite question (kappa = 0.71). However, the presence of nodular densities (kappa = 0.56) and parenchymal consolidation (kappa = 0.57) had moderate agreement and agreement for lung volume was low (25). Compare these reliability coefficients to the test–retest reliability or measures of internal consistency reliability of commonly used patients’ self-report measures such as the European Organization for Research and Treatment of Cancer Core Quality of Life Questionnaire (EORTC QLQ-C30) or the FACT, both well-validated instruments used by the RTOG and other CTCG. For example, evaluation of EORTC QLQ-C30 reliability measured by Pearson’s correlation coefficient was high for all functional subscales, with a range from .82 for cognitive and role function to .91, and for physical function and for the symptom scales (i.e., nausea/vomiting, fatigue, and pain), the coefficients were .63, .83, and .86, respectively (26). Test–retest for the FACT-General total score has been documented as .92 and from .82 to .88 for the various subscales while internal consistency for the total scale was documented as .89 (27).

Of course psychometric validity and reliability are only some of the factors to consider in choosing outcome measures. It is paramount that the research question guides the choice of the measures selected and that good science and common sense prevail. While psychometrically sound measures are critical for the primary endpoint, if validated measures are not available to capture the variables of interest, then exploratory aims using newly developed measures may be reasonable.

It is the identification of clinical trials appropriate for endpoints identified in the Model and the guidance and facilitation of the choice of appropriate outcome measures that is the unique role of the HSRO Committee in RTOG. Studies guided by the RTOG Outcomes Model provide supplementary nontraditional CTCG assessments and outcomes including quality-adjusted life years and a cost–utility analyses. In the examples presented, these outcomes have served to better quantify the benefit of one therapy over alternative therapies as in the example of multimodality therapy for lung cancer and to add evidence to clinical outcomes when clinical outcomes alone have not been strong enough to change clinical practice as in the example of palliative radiotherapy for painful bone metastasis.

The success of the RTOG Outcomes Model in guiding the choice of endpoints in particular trials has led to wide acceptance in the group. However, the question has been raised as to whether all RTOG trials are guided by the Model. The answer is no. The reason is twofold: not all trials would benefit from this framework (i.e., phase I and some of the molecular epidemiology and translational research studies) and resources are finite. The reality of limited resources has directed strategies to prioritize the use of the Model, which include focusing on phase III trials and the search for external funding (i.e., NIH, Foundation and pharmaceutical support).

**Conclusion**

RTOG has evolved from the assessment of clinical outcomes and QOL to a more comprehensive model that includes a triad of outcomes (clinical, humanistic, and economic) identified as important endpoints in radiation therapy clinical treatment and symptom management trials. RTOG is seeking to use the information gained from the model to guide more targeted symptom interventional studies and to assess improvements in PROs, patient preferences, and economic outcomes. In RTOG, PROs are considered complementary to clinician-reported assessments of toxicity. In point of fact, data most important to the patient, and decision making of future patients, are most likely missing from
clinical trials that only use medical professionals’ ratings of CTC data and that do not include patients’ self-reports of particularly subjective symptoms. The unique contribution to the RTOG of the HSRO Committee is the selection and use of PRO measures to give “voice” to the patient in the study of clinical trials outcomes including symptoms.

References