

Evaluating Health-Related Quality of Life in Cancer Clinical Trials: The National Cancer Institute of Canada Clinical Trials Group Experience

David Osoba, MD,¹ Andrea Bezjak, MDCM,² Michael Brundage, MD, PhD,³ Joseph Pater, MD⁴

¹QOL Consulting, West Vancouver, BC, Canada; ²Department of Radiation Oncology, Princess Margaret Hospital, University of Toronto, Toronto, ON, Canada; ³Radiation Oncology, Kingston Regional Cancer Center, Queens University, Kingston, ON, Canada; ⁴NCIC Clinical Trials Group, Queens University, Kingston, ON, Canada

ABSTRACT

Introduction: The National Cancer Institute of Canada (NCIC) Clinical Trials Group (CTG) Quality of Life (QOL) Committee was initiated in 1986.

Purpose: The purpose of this review is to describe the evolution of the Committee's work and to highlight key developments such as the formulation of a policy regarding health-related quality-of-life (HRQOL) assessment, the provision of guidelines to ensure completion of HRQOL data within the protocol requirements, the rationale behind the choice of HRQOL instruments, the timing of assessments and the development of data analytic methods. These

developments are illustrated with examples from CTG studies.

Recommendations: There is a lack of concordance between conventional toxicity data and HRQOL data and comparative studies designed to elucidate these differences are to be encouraged. Also, more studies are required to compare different analytic strategies and to determine how much missing data is acceptable, particularly in oncology studies where attrition is inevitable.

Keywords: assessment, clinical trials, oncology, QOL, quality of life.

The Quality of Life Committee—Description

The experience of the National Cancer Institute of Canada (NCIC) Clinical Trials Group (CTG) in health-related quality-of-life (HRQOL) assessment began in 1986 with the formation of a Quality of Life (QOL) working group. The role of the working group was to provide educational opportunities to the CTG clinical investigators and other CTG personnel. A standing QOL Committee was established in 1987 [1]. Subsequently, HRQOL assessment in the CTG has evolved gradually in keeping with developing worldwide knowledge. In several ways, the CTG has been instrumental in introducing new information during this evolution. Some of these advances are detailed in this article, which gives particular emphasis to how our policy and procedures have led to useful methods and results, and how this experience may be useful for developing Food and Drug Administration guidance for labeling claims.

Policy

The Committee developed a policy pertaining to HRQOL assessment in the context of clinical trials [1].

Address correspondence to: David Osoba, QOL Consulting, 4939 Edendale Court, West Vancouver, BC, Canada V7W 3H7. E-mail: david_osoba@telus.net
10.1111/j.1524-4733.2007.00278.x

The policy stated that “there should be a statement about the anticipated impact on quality of life with every proposed Phase III clinical trial and whether QOL measures will be incorporated in the protocol.”

As a result of this policy, almost all of the 71 trials initiated by the CTG since its adoption in 1987 include HRQOL components. This provides the CTG with extensive experience to determine which procedures have worked best in given situations, and the results have provided new information. Between 1992 and 2006, the QOL Committee published 45 full articles, 69 abstracts, and 6 miscellaneous publications. The presence of a CTG policy, however, does not always lead to the inclusion of HRQOL assessments in intergroup trials if the lead group does not feel it necessary to include such assessment.

Writing Guidelines

The QOL Committee provides writing guidelines for the inclusion of HRQOL assessment in clinical trials protocols [1,2]. HRQOL components should be a part of the main protocol and not an add-on. Inclusion of HRQOL components in the main protocol is desirable because adding on separate protocols creates the impression that HRQOL assessment is either not very important or that it is an afterthought.

The writing guidelines provide explicit instructions for the sections of the protocol dealing with the

introduction, rationale and hypothesis, objectives of HRQOL assessment, eligibility criteria, study design, sample size, instrument description, instructions for administration of the instruments, timing of assessments [3], analysis, and wording of the consent form.

Instructions to Clinical Research Associates (CRAs)

At the outset, the QOL Committee developed instructions for the integration of HRQOL measurement into Phase III clinical trials [1]. These instructions are used by CRAs, office data managers, and others who will be involved in collecting the data. They include the rationale for the HRQOL component, how to present HRQOL questionnaires to patients, how to collect them after completion, and how to transmit the completed questionnaires to the central CTG office. As a result, all personnel involved with an HRQOL component in a trial understand the need for measuring it and the procedure for standardized collection of the data. Because central office staff who conduct the main part of the clinical trial are the same individuals who receive the HRQOL data, there is integration of all components of the trial by one group of personnel. We believe that this is an advantage over having separate personnel dealing with the HRQOL component and with the primary component of the trial (i.e., the primary objective).

A particularly important instruction to the CRAs is that they should obtain a completed HRQOL questionnaire on a patient before calling the central office for randomization instructions. This is intended to make certain that questionnaire completion rates at baseline are high. Nevertheless, if a patient is not able to answer a questionnaire because of language barriers, but meets all the other inclusion criteria, the patient is not excluded from participation in the study.

Liaison with Disease Site Groups (DSGs)

The QOL Committee members provide liaison to DSGs and educational resources to members of the CTG. Members of the QOL Committee act as liaison members between the QOL Committee and specific DSGs so that the HRQOL component of a proposed trial can be incorporated into the protocol at an early stage of the planning process. In addition, feedback to the DSGs on the success of collection of the HRQOL data can be provided as the trial progresses. To ensure that DSGs are well acquainted with the need to collect QOL data and what this entails, the QOL liaison person is a member of the executive committee for each DSG.

Completion Rates of HRQOL Instruments (Compliance)

The above policy and procedural instructions are intended to produce high questionnaire completion rates so as to keep avoidable (random) missing data to a minimum. Completion rates are determined as

follows: number of questionnaires received with sufficient items answered to be deemed complete [4]; the number of questionnaires received over the number of patients enrolled in the trial; and the number of questionnaires received over the number expected (number of patients still on study and required to complete questionnaires according to the protocol). To date, almost all trials with HRQOL components have baseline completion rates higher than 97% [5–7]. On-study completion rates are lower, but the number completed over the number that can be expected to be completed is usually more than 80% (unpublished data) [5–7]. These completion rates are among the best in the world, and we believe that they are a consequence of the diligent efforts of our CRAs and central office personnel.

Choice of HRQOL Instruments

When the QOL Committee was formed, there were only a small number of HRQOL instruments from which to choose. Only two or three had been developed for use in patients with cancer. One of the members of the QOL Committee became a liaison member from the CTG to the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Study Group and was made aware of the properties of the Quality of Life Questionnaire (QLQ) being developed by that group [8]. The details of the questionnaire's reliability and validity were reviewed by the CTG QOL Committee, and it was decided to adopt the use of this questionnaire for CTG purposes. The decision was made to use this questionnaire as our standard questionnaire in clinical trials for a number of reasons. First, it provides separate domain and single-item scores rather than a single aggregated or summary score. Having separate scores allows a detailed picture of how different domains of HRQOL are affected by a disease and its treatment. Although a single aggregated score is appealing because it is simpler to deal with, it hides changes in HRQOL domains that move in opposite directions or that may differ according to the treatment a patient receives. Second, using the same instrument in a variety of disease sites allows us to become very familiar with its properties and behavior in a variety of cancer populations. Third, it allows us to make comparisons across clinical trials at different disease sites if we wish to do so. Finally, the use of one questionnaire allows for simplicity of administration at the clinic level and in data management at the central office. Eventually, we began to use other questionnaires if we were participating in a trial initiated by another cooperative clinical trials group or if it was felt that another instrument had some properties that made it desirable for use in a particular trial. We have used 24 different instruments, but the EORTC QLQ-C30 (the core instrument with 30 items) has been used in 51 of 71 (71.8%) trials.

Development of Study-Specific Checklists

Generic and general QOL instruments need to be supplemented with disease-specific modules [9] or with study-specific items to provide a more complete collection of the necessary HRQOL data [10]. Almost no disease-specific modules/checklists existed for use in cancer populations by the early 1990s. By study-specific items, we refer to items that are specific to a particular trial and that ask questions not covered in the core questionnaire or a disease-specific module. Examples include symptoms caused by the disease or its treatment (e.g., difficulty passing urine in prostate cancer, bleeding from the vagina, the bladder, or rectum after radiation therapy (RT) for cervical cancer or for prostate cancer, or fever after treatment with interferon or from infection after chemotherapy.) Study-specific checklists are supplementary to disease-specific modules and general or generic instruments. Because they consist of single items, they cannot be tested for reliability or validity in the usual ways, but they can be assessed for understandability and acceptability to the patient.

Several study-specific checklists were developed by the QOL Committee. As an example, one of these has been used in the MA.14 breast cancer trial of tamoxifen versus tamoxifen plus octreotide in postmenopausal women with primary breast cancer who have undergone definitive surgical management [11]. The analysis of the MA.14 HRQOL checklist results shows that, in the short term, the selective checklist item scores discriminate between patients differing in clinical status and are responsive to change over time as well as to treatment-induced differences. The selective checklist item scores correlate statistically significantly with the patient's global QOL scores and provide additional information to that provided by the EORTC QLQ-C30 core questionnaire and the EORTC breast cancer module QLQ-BR23 (<http://www.cancer.gov/clinicaltrials/EORTC-15931>).

Results

In this section, the results of several CTG studies are presented to illustrate how HRQOL assessment helps to better understand the effects of cancer and its treatment on patients' lives, to determine which treatment produces better HRQOL in randomized trials, and to provide recommendations that may help in the formulation of a guidance document. Many of our findings are based on large numbers of patients pooled together, in some cases, from several trials. It was possible to do this because of our early decision to use the same instrument (the QLQ-C30) in most of our trials. Had we used different instruments in each trial, the numbers could not have been combined and would have been too small to provide reliable data.

Validation and Reliability Studies

Further validation (reliability and validity) of the QLQ-C30 was established in ovarian, breast, and lung cancer [12]. In addition, CTG studies helped to modify the original version of the QLQ-C30 into the second and third versions of the questionnaire [13,14].

Nausea and Vomiting after Chemotherapy

Studies of chemotherapy-induced nausea and vomiting in 832 chemotherapy-naive patients, enrolled in studies of either highly or moderately emetogenic chemotherapy, elucidated the deleterious effects of nausea and vomiting together and of each independently on physical, social, and cognitive functioning as well as on fatigue, appetite loss, and global QOL [15,16]. A study of risk factors associated with postchemotherapy nausea and vomiting in these patients revealed that low social functioning, prechemotherapy nausea, being female, increased fatigue, and a lower performance status were associated with a higher risk of either vomiting or nausea after highly emetogenic chemotherapy [17]. After moderately emetogenic chemotherapy, however, some prognostic factors seemed to be inconsistent, and the probability of postchemotherapy nausea and vomiting was more strongly influenced by the type of chemotherapy given and the type of antiemetic used rather than by patient or by environmental factors [18].

It was also established in a study of moderately emetogenic chemotherapy that patients could be relied upon to pay attention to the time frame of the questions, because their answers were consistent with the expectation that postchemotherapy nausea and vomiting would have a greater impact on HRQOL during the first 3 days after chemotherapy than during the subsequent 4 days [19]. Very short time frames (e.g., 1 day) provide inconsistent data because the time frame is too short, and therefore 3- to 7-day time frames are recommended to provide the most reliable data [20].

Fatigue

In 2390 patients with a variety of cancers, greater fatigue severity was found to be associated with being female, metastatic disease, and poorer performance status. The oldest patients and patients with breast cancer reported less fatigue, while patients with ovarian cancer or lung cancer had greater fatigue [21]. Patients in antiemetic trials whose nausea and vomiting were well controlled showed less fatigue than did those with poor control of nausea and vomiting. Nevertheless, patients who had complete control of nausea and vomiting after highly emetogenic chemotherapy still had greater fatigue than at baseline [16]. Thus, not all postchemotherapy fatigue can be attributed to poor

control of nausea and vomiting. These studies were important in establishing the concept that complete control of nausea and vomiting following chemotherapy is ideal [22].

Prognostic Factors for Survival and Treatment Effects

A study of prognostic factors for survival in a heterogeneous group of 474 patients who had either localized or metastatic disease and who had been entered into chemotherapy trials showed that lower global QOL scores on the QLQ-C30 at baseline were independently associated with shorter survival [23].

In other clinical trials, CTG studies have shown that pretreatment global QOL scores were a predictor of on-treatment global QOL in patients with malignant melanoma [24]. In advanced ovarian cancer, patients treated with cyclophosphamide and cisplatin had less deleterious impact on HRQOL during treatment than patients treated with paclitaxel and cisplatin. The latter patients reported more difficulty with myalgia and neurosensory problems than did the former, but by a year after treatment, both groups were similar [25].

An intergroup trial between an Italian group and the CTG in Stage II and IV nonsmall-cell lung cancer (NSCLC) found that although patients treated with gemcitabine plus vinorelbine reported fewer problems with appetite, vomiting, alopecia, and ototoxicity than patients receiving cisplatin-based chemotherapy, their lung cancer symptoms were not as well controlled and there was no overall difference in global QOL between the cisplatin and non-cisplatin-based chemotherapy [26]. QOL was the primary study end point in that study, and provided clinically useful information that was not previously known. A large study of the targeted agent, erlotinib, versus placebo as second- or third-line treatment for Stage IV NSCLC revealed a clinically and statistically significantly longer median time to deterioration of predefined index symptoms for patients on the erlotinib arm (e.g., 4.7 vs. 2.9 months for dyspnea) as well as improvements in global QOL and physical function. These findings confirmed that the observed survival benefit is truly of clinical benefit as it provides better palliation and QOL [27].

Another study in extensive-stage small-cell lung cancer, however, showed how HRQOL can be adversely affected by a new, dose-intensive therapy as compared with standard therapy [28]. This trial was stopped before completion of accrual because of an excess of early deaths, but the HRQOL data also showed a clear worsening of physical functioning, fatigue, and global QOL in the dose-intensive arm. A randomized trial of preoperative versus postoperative RT in 190 patients with extremity soft tissue sarcoma showed that those who received postoperative

RT had better limb function and less pain at 6 weeks after surgery than did patients with preoperative RT [29]. There were no differences at later time points. The conclusion was that timing of RT had a minimal impact on function in the first year after surgery.

Observer Ratings versus Direct Patient-Reported Outcomes

A study using 12 simulated patients asked seven experienced clinical data managers to score toxicity grades using the NCIC CTG and the World Health Organization (WHO) toxicity scales [30]. Modest levels of inter-rater reliability were demonstrated with kappa values that ranged from 0.50 to 1.00 in laboratory-based categories and from -0.04 to 0.82 for clinically based categories. Proportions of agreement for clinical categories ranged from 0.52 to 0.98. Condensing the toxicity grades improved statistics of agreement, but substantial lack of agreement remained (kappa range -0.04 to 0.82; proportions of agreement range 0.67-0.98). Thus, experienced data managers, when interviewing patients, draw varying conclusions regarding toxic effects experienced by such patients. Neither the NCIC CTG-expanded toxicity scale nor the WHO standard toxicity scale demonstrated a clear superiority in reliability, although the breadth of toxic effects recorded differed.

In addition, a low correlation exists between patient-reported toxicity and toxicity grade by observers [31]. Between 1988 and 1990, the NCIC CTG conducted a multicenter Phase III trial that compared two immunomodulating agents as long-term adjuvant therapy in resectable malignant melanoma [24]. Incidence and severity data on 11 symptoms of particular interest were collected by each of three methods: the case report flow sheet (FS) completed by study personnel, the Symptom Diary (SD), and the QLQ-C30 (the latter two patient-reported). Both the FS and SD included a preset list so that each toxicity required a graded response (0 = none).

Ten of the 89 available cases were randomly selected for an initial analysis. In four patients, all three methods produced either identical or only slightly differing records of the expected toxicities. In five patients there were discrepancies. In one patient, each method documented different toxicities. In four cases, the QLQ-C30 picked up more symptoms than the SD or FS. The SD in turn picked up more symptoms than did the FS. Finally, one case was not evaluable as treatment was stopped after 2 weeks, but the QLQ-C30 was not done until 16 weeks. Preliminary results thus indicate that reporting of treatment-related symptoms can differ substantially among these three methods of data collection.

A more recent study involving 303 patients with advanced breast cancer compared the level of agreement between patients' self-evaluations and health-care personnel's accounts of patients' symptoms [32]. Fifteen "matched" symptoms from the QLQ-C30 and the NCIC CTG-expanded Common Toxicity Criteria were evaluated over seven common time points. Agreement was only fair to slight (kappas from 0.012 to 0.378) between patients and health care personnel and worsened over time. In general, patients reported many more symptoms than did health care personnel. Therefore, the use of only toxicity grading by health-care personnel may result in under-reporting of toxicity, thus altering study results. Further studies are required to explain these discrepancies.

Analysis of HRQOL Data

The analysis and reporting of HRQOL data has undergone a gradual evolution. Initially, NCIC investigators sought to determine whether there was a statistically significant difference in the means of domain mean scores between the treatment arms of a study [24]. Those with higher scores may be healthier than those with lower scores. Using only mean scores may result in a possible survivor bias.

Subsequently, several studies have compared the means of the changes in HRQOL scores from baseline between treatment groups [16–18]. In this approach, the change in a patient's scores is calculated by subtracting the baseline score from subsequent on-study scores. Thus, only scores of patients who are still on study are used. The advantage is that a spurious increase in the group mean scores is eliminated because an individual patient's on-study scores are compared only with that patient's own baseline scores. The disadvantage, however, is that analysis of this subset of patients may limit the generalizability of the results to the whole sample. Also, when the individual differences are taken together, the mean change in scores from baseline between groups does not inform us of the clinical significance of the change in scores. For example, as alluded to earlier, is a difference of 5% in the mean change scores clinically meaningful? Knowledge of the change for the entire group also does not inform us of the differences reported by individual patients, some of whom may have experienced larger changes in scores while others experienced smaller changes.

Measuring Clinically Meaningful Change

Is a change of 3 or 5 points on a scale of 0–100 important from a clinical perspective? The CTG chose to study this question by building on an approach used to determine the minimal important difference (MID) as suggested by the studies of a group at

McMaster University [33,34]. A Subjective Significance Questionnaire was used to determine the smallest amount of change on four QLQ-C30 domains that was perceptible to patients with advanced breast cancer and small-cell lung cancer as a change from the previous time that they had completed the QLQ-C30. It was found that the smallest subjectively significant change perceptible to patients is between 5% and 10% of the breadth of the QLQ-C30 for physical, social, and emotional functioning and global QOL [35]. A change of 11% to 20% is moderate and changes of more than 20% are large. On average, these changes correspond to effect sizes of less than 0.5, 0.5 to 0.8, and more than 0.8, respectively [35].

Several other studies in patients with different diseases, completing different questionnaires, have shown that a change of about 7% to 8% of the scale breadth is perceived as a change from baseline [33–36]. By using other external anchors to determine a MID, a similar difference has been determined [37,38]. The magnitude of this difference is about 0.4 of a standard deviation or of an effect size [39,40]. Thus, this difference can be used as a cut point to distinguish patients who have experienced an HRQOL benefit from those who have not experienced a benefit. Using a cut point right at the MID, however, may include "false positives" as having experienced a true benefit. Therefore, a higher cut point (i.e., 10%) to distinguish those who have had a true HRQOL benefit from those who have not is recommended [41–47]. A 10% cut point is about 0.5 of a standard deviation or of an effect size [35,39,40]. In addition to being used for determining the proportions of patients with an HRQOL benefit, the cut point can also be used to compare the duration of benefit within and between groups of patients. Also, the magnitude of these changes can be helpful for calculating sample sizes required to detect a specified change in clinical trials.

NCIC CTG Recommendations for Analysis

The NCIC CTG recommends a relatively simple, clinically practical analysis that consists of four steps [47]. After calculation of the questionnaire completion rates and the baseline scores, the individual change scores from baseline are determined. Then, the proportions of patients who have reported a predetermined clinically significant change score (usually more than 10%) are calculated, and the differences in the number of patients who have benefited in each treatment group are tested for statistical significance. An advantage of this approach of calculating the proportion of patients with an "HRQOL response" is that patients with missing data may still be included in an intention-to-treat analysis.

A recent study compared the NCIC CTG QOL standard protocol with two other methods, summary measure and summary statistic-based approaches, and linear mixed model-based methods, in analyzing HRQOL data based from a randomized clinical trial in patients with advanced breast cancer [48]. The different statistical approaches have both advantages and disadvantages. The CTG standard protocol is easy to implement and takes clinical importance into account in the analysis of the QOL data. Other summary measures and statistics-based approaches are also simple to interpret. The time effects cannot be assessed with these approaches. Model-based methods take the correlation between repeated measurements into account and could test the treatment effect over time. For a comprehensive exploratory analysis, model-based methods would seem to be essential. The assumptions underlying the model-based approaches, however, are difficult to verify and have a strong impact on the validity of the analysis results.

This study did not identify a method that is better than all other approaches, and our suggestion is that, in the analysis of QOL data, different methods should be explored to assess the robustness of the results. Other analyses, such as growth curves [28,49] or pattern mixture models [50,51], may be applied. An additional useful analysis is to calculate the proportion of patients reporting a “HRQOL response” within a treatment group. This is useful not only in nonrandomized Phase II trials, where there is no comparison group [45], but also in Phase III randomized studies [44,46].

Conclusions

The strategies for and the approach to assessing HRQOL in clinical trials have evolved over time within the CTG. Some of the strategies that were set in place at the outset, such as developing a policy, producing writing guidelines and explicit instructions to CRAs and other personnel, and assigning liaison members to DSGs, have proved useful over time. These strategies have produced high questionnaire completion rates with a minimum of randomly missing data due to administrative error. A rationale has been presented for our decision to use a single questionnaire, the QLQ-C30, in most of our trials. The QOL Committee of the CTG has produced simple guidelines for the analysis of HRQOL data that culminate in a determination of the proportions of patients who benefit from treatment. Whether our simple response-based analysis is sufficiently robust for most purposes is still uncertain, and it needs to be tested by comparison with more complicated analytic approaches.

Key points and recommendations from the NCIC CTG experience

- 1 It is possible to collect quality data, even when HRQOL outcomes are complex
 - a. Institute a clear policy mandating HRQOL assessments in appropriate Phase III trials
 - b. Provide guidelines on the required HRQOL content in clinical trial protocols
 - c. Establish methods to minimize randomly missing HRQOL scores
- 2 Conventional toxicity data assessment has limited reliability for subjective end points
 - a. Validity and reliability of subjective (non-laboratory-based) data are poor
 - b. There is a lack of concordance between toxicity data and HRQOL data
 - c. Encourage comparative studies designed to elucidate these differences
- 3 QOL scores have repeatedly been validated and are subject to design detail
 - a. Validated as independent measures of outcome
 - b. Validated as independent predictors of outcome
 - c. Sensitive to time frame of interest
- 4 Simple analyses are preferred
 - a. Encourage comparative studies of different analysis strategies
 - b. Encourage studies to explore how much missing data is “acceptable”
- 5 Single-item outcome assessments should not be ignored and may have validity
 - a. Within a validated questionnaire
 - b. As single-item check lists
 - c. Encourage studies to determine the value of single items

HRQOL, health-related quality of life; NCIC CTG, National Cancer Institute of Canada Clinical Trials Group.

Financial support: None.

References

- 1 Osoba D. The quality of Life Committee of the Clinical Trials Group of the National Cancer Institute of Canada: organization and functions. *Qual Life Res* 1992;1:211–8.
- 2 Osoba D. Guidelines for measuring health-related quality of life in clinical trials. In: Staquet MJ, Hays RD, Fayers PM, eds. *Quality of Life Assessment in Clinical Trials*. Oxford: Oxford University Press, 1998.
- 3 Osoba D. Rationale for the timing of health-related quality-of-life (HQL) assessments in oncological palliative therapy. *Cancer Treat Rev* 1996;22 (Suppl. A):69–73.
- 4 Fayers P, Aaronson N, Bjordal K, Sullivan M. *EORTC QLQ-C30 Scoring Manual*. Brussels, Belgium: EORTC Study Group on Quality of Life, 1995.
- 5 Sadura A, Pater J, Osoba D, et al. Quality-of-life assessment: patient compliance with questionnaire completion. *J Natl Cancer Inst* 1992;84:1023–6.
- 6 Osoba D, Dancy J, Zee B, et al. Health-related quality of life studies of the National Cancer Institute of Canada Clinical Trials Group. *J Natl Cancer Inst Monogr* 1996;20:107–11.
- 7 Osoba D, Zee B. Completion rates in health-related quality-of-life assessment: approach of the National Cancer Institute of Canada Clinical Trials Group. *Stat Med* 1998;17:603–12.

- 8 Aaronson NK, Ahmedzai S, Bullinger M, et al. The EORTC core quality of life questionnaire: interim results of an international field study. In: Osoba D, ed. *Effect of Cancer on Quality of Life*. Boca Raton, FL: CRC Press, 1991.
- 9 Aaronson NK, Bullinger M, Ahmedzai S. A modular approach to quality-of-life assessment in cancer clinical trials. In: Scheurlen H, Kay R, Baum M, eds. *Cancer Clinical Trials: A Critical Appraisal. Recent Results in Cancer Research, Vol. III*. Berlin: Springer-Verlag, 1988.
- 10 Osoba D. Self-rating symptom checklists: a simple method for recording and evaluating symptom control in oncology. *Cancer Treat Rev* 1993;19:43-51.
- 11 Chen D. Supplementary health-related quality of life assessment in cancer clinical trials: added value to standard HRQoL instruments? Master's Thesis, Queen's University, Kingston, Ontario, Canada; September 2005.
- 12 Osoba D, Zee B, Pater J, et al. Psychometric properties and responsiveness of the EORTC Quality of Life Questionnaire QLQ-C30 in patients with breast, ovarian and lung cancer. *Qual Life Res* 1994;3:353-64.
- 13 Fayers P, Aaronson N, Bjordal K, et al. *EORTC Scoring Manual, 2nd edn*. Brussels, Belgium: Quality of Life Unit, EORTC Data Centre, 1999.
- 14 Fayers PM, Aaronson N, Bjordal K, et al. *EORTC QLQ-C30 Scoring Manual, 3rd edn*. Brussels, Belgium: EORTC Quality of Life Group, 2001.
- 15 Osoba D, Zee B, Warr D, et al. Quality of life studies in chemotherapy-induced emesis. *Oncology* 1996;53 (Suppl. 1):92-5.
- 16 Osoba D, Zee B, Warr D, et al. Effect of postchemotherapy nausea and vomiting on health-related quality of life. *Support Care Cancer* 1997;5:307-13.
- 17 Osoba D, Zee B, Pater J, et al. Determinants of postchemotherapy nausea and vomiting in patients with cancer. *J Clin Oncol* 1997;15:116-23.
- 18 Pater J, Slamet L, Zee B, et al. Inconsistency of prognostic factors for postchemotherapy nausea and vomiting. *Support Care Cancer* 1994;2:161-6.
- 19 Pater J, Osoba D, Zee B, et al. Effects of altering the time of administration and the timeframe of quality of life assessments in clinical trials: an example using the EORTC QLQ-C30 in large antiemetic trials. *Qual Life Res* 1998;7:273-8.
- 20 Huisman SJ, van Dam FSAM, Aaronson NK, Hanewold GJFP. On measuring complaints of cancer patients: some remarks on the time span of the question. In: Aaronson NK, Beckman J, eds. *The Quality of Life of Cancer Patients*. New York: Raven Press, 1987.
- 21 Pater JL, Zee B, Palmer M, et al. Fatigue in patients with cancer: results with National Cancer Institute of Canada Clinical Trials Group studies employing the EORTC QLQ-C30. *Support Care Cancer* 1997; 5:410-4.
- 22 Gralla RJ, Osoba D, Kris MG, et al. Recommendations for the use of antiemetics: evidence-based, clinical practice guidelines. *J Clin Oncol* 1999;17:2971-94.
- 23 Dancy J, Zee B, Osoba D, et al. Quality of life score: an independent prognostic variable in a general population of cancer patients receiving chemotherapy. *Qual Life Res* 1997;6:151-8.
- 24 Osoba D, Zee B, Sadura A, et al. Measurement of quality of life in an adjuvant trial of gamma interferon versus levamisole in malignant melanoma. In: Salmon SE, ed. *Adjuvant Therapy of Cancer VII*. Philadelphia: JB Lippincott Co., 1993: 412-6.
- 25 Bezjak A, Tu D, Bacon M, et al. Quality of life in ovarian cancer patients: comparison of paclitaxel plus cisplatin, with cyclophosphamide plus cisplatin in a randomized study. *J Clin Oncol* 2004;22:4595-603.
- 26 Gridelli C, Gallo C, Shepherd FA, et al. Gemcitabine plus vinorelbine compared with cisplatin plus vinorelbine or cisplatin plus gemcitabine for advanced non-small-cell lung cancer: a Phase III trial of the Italian GEMVIN Investigators and the National Institute of Canada Clinical Trials Group. *J Clin Oncol* 2003;21:3025-34.
- 27 Shepherd F, Pereira J, Ciuleanu T, et al. Erlotinib in previously treated non-small-cell lung cancer. *N Engl J Med* 2005;353:123-32.
- 28 Murray N, Livingston RB, Shepherd FA, et al. Randomized study of CODE versus alternating CAV/EP for extensive-stage small-cell lung cancer: an intergroup study of the National Cancer Institute of Canada Clinical Trials Group and the Southwest Oncology Group. *J Clin Oncol* 1999;17:2300-8.
- 29 Davis AM, O'Sullivan B, Bell RS, et al. Function and health status outcomes in a randomized trial comparing preoperative and postoperative radiotherapy in extremity soft tissue sarcoma. *J Clin Oncol* 2002;20:4472-7.
- 30 Brundage MD, Pater JL, Zee B. Assessing the reliability of two toxicity scales: implications for interpreting toxicity data. *J Natl Cancer Inst* 1993;85:1138-48.
- 31 Paul N, Pater J, Whitehead M, Sadura A. Methods of toxicity data collection: an evaluation of the relative effectiveness of the case report flow sheet, the patient symptom diary, and the quality of life questionnaire. *Control Clin Trials* 1991;12:648.
- 32 Savage C, Pater J, Tu D, Norris B. He said/she said: how much agreement is there on symptoms between common toxicity criteria and quality of life? *Proc Am Soc Clin Oncol* 2002;21:386a (abstract 1540).
- 33 Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989; 10:407-15.
- 34 Juniper ED, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. *J Clin Epidemiol* 1994;47:81-7.
- 35 Osoba D, Rodrigues G, Myles J, et al. Interpreting the significance of changes in health-related quality-of-life-scores. *J Clin Oncol* 1998;16:139-44.
- 36 Ehrlich EW, Davies GM, Watson DJ, et al. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis

- index questionnaire and global assessments in patients with osteoarthritis. *J Rheumatol* 2000;27:2635–41.
- 37 King MT. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. *Qual Life Res* 1996;5:555–67.
- 38 Cella D, Eton DT, Fairclough DL, et al. What is a clinically meaningful change on the Functional Assessment of Cancer Therapy-Lung (FACT-L) questionnaire? Results from Eastern Cooperative Oncology Group (ECOG) Study 5592. *J Clin Epidemiol* 2002; 55:285–95.
- 39 Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003;41:582–92.
- 40 Farivar SS, Kiu H, Hays RD. Another look at the half standard deviation estimate of the minimally important difference in health-related quality of life scores. *Expert Rev Pharmacoeconomics Outcomes Res* 2004; 4:521–9.
- 41 Osoba D, Northfelt DW, Budd DW, Himmelberger D. Effect of treatment on health-related quality of life in AIDS-related Kaposi's sarcoma: a randomized trial of pegylated liposomal doxorubicin versus doxorubicin, bleomycin and vincristine. *Cancer Invest* 2001; 19:573–80.
- 42 Osoba D, Tannock IF, Ernst DS, Neville AJ. Health-related quality of life in men with metastatic prostate cancer treated with prednisone alone or mitoxantrone and prednisone. *J Clin Oncol* 1999; 17:1654–63.
- 43 Osoba D. A taxonomy of the uses of health-related quality of life (HRQL) instruments in cancer care and the clinical meaningfulness of the results. *Med Care* 2002;40(Suppl.):III31–8.
- 44 Osoba D, Slamon DJ, Burchmore M, Murphy M. Effects on quality of life of combined trastuzumab and chemotherapy in women with metastatic breast cancer. *J Clin Oncol* 2002;20:3106–13.
- 45 Osoba D, Brada M, Yung WKA, Prados MD. Health-related quality of life in patients with anaplastic astrocytoma during treatment with temozolomide. *Eur J Cancer* 2000;36:1788–95.
- 46 Osoba D, Brada M, Yung WKA, Prados M. Health-related quality of life in patients treated with temozolomide versus procarbazine for recurrent glioblastoma multiforme. *J Clin Oncol* 2000;18:1481–91.
- 47 Osoba D, Bezjak A, Brundage M, et al. Analysis and interpretation of health-related quality-of-life data from clinical trials: basic approach of The National Cancer Institute of Canada Clinical Trials Group. *Eur J Cancer* 2005;41:280–7.
- 48 Liu J. Analysis of quality of life data from a clinical trial in patients with advanced breast cancer with a comparison of different statistical approaches. Master's Thesis, Queen's University, Kingston, Ontario, Canada; August 2002.
- 49 Zee B. Growth curve model analysis for quality of life data. *Stat Med* 1998;17:757–66.
- 50 Fairclough D, Peterson H, Cella D, Bonomi P. Comparison of several model-based methods for analyzing incomplete quality of life data in clinical trials. *Stat Med* 1998;17:781–96.
- 51 Fayers P, Machin D. *Quality of Life: Assessment, Analysis and Interpretation*. Chichester: John Wiley & Sons, 2000.