Quality of life and/or symptom control in randomized clinical trials for patients with advanced cancer

F. Joly1, J. Vardy2, M. Pintilie3 & I. F. Tannock2*

1Department of Medical Oncology, Centre François Baclesse, Caen, France; 2Department of Medical Oncology and Hematology; 3Department of Biostatistics, Princess Margaret Hospital and University of Toronto, Toronto, Canada

Received 4 July 2006; revised 7 November 2006; revised 13 December 2006; accepted 8 March 2007

Background: Measures reflecting quality of life (QoL) or symptom control should be included as major endpoints in most phase III trials for patients with advanced cancer. Here we review the use of such endpoints.

Methods: We evaluated methodological aspects relating to QoL or symptom control in randomized controlled trials (RCTs) that included ≥150 patients, published from 1994 to 2004, using a 10-point checklist.

Results: Of 112 RCTs that met our criteria, few were rated as high quality: 22% defined QoL or symptom control as a primary endpoint; 19% established an a priori hypothesis relevant to palliation and 21% defined minimal differences in QoL or symptom scores that were clinically meaningful. Most trials (81%) analyzed differences between mean or median scores across groups and only 21% defined the proportion of individual patients who met criteria for palliative response. Only 15% of the studies met more than 5/10 criteria from our checklist. There was improvement over time in methodology and reporting.

Conclusions: Current standards for analyzing QoL and symptom control in RCTs are poor. Definition of a palliative endpoint, with an a priori hypothesis, is essential; defining the proportion of patients with palliative response is preferred. The proposed checklist could raise standards of reporting in future RCTs.

Key words: quality of life, randomized controlled trial, symptom control

introduction

Many clinical trials compare treatments for patients with advanced metastatic cancer: the therapies do not offer a meaningful chance of cure, and the aim of treatment is palliation, which requires improvement in either the duration or quality of survival (or both). Palliation can refer to either of two approaches: (i) pain and symptom management as described by the World Health Organization (palliative care), and/or (ii) non-curative treatments such as chemotherapy, radiotherapy and surgery (palliative oncological treatment) [1]. In this review, we focus on the second approach.

When treatment is given with palliative intent, it is essential to assess outcome with palliative endpoints [2]. Overall survival is always an important endpoint in such trials, but only a few treatments have led to significant improvements in duration of survival for common metastatic tumors of adults [3]. Improvement in symptoms and in quality of life (QoL) is often a more feasible goal [4]. During the past decade there has been increasing recognition of the importance of integrating measures of QoL and/or symptom control with other indicators of efficacy in randomized clinical trials (RCTs) [4–6]. This has been stimulated not only by growing recognition that many cancer treatments may have a higher probability of improving the quality than the duration of survival, but also by the recognition of QoL or symptom control endpoints by agencies such as the United States Food and Drug Administration (FDA) for the registration of new anticancer treatments [7, 8]. Although these outcomes are being included more frequently in RCTs, they are seldom applied with the same rigor as the more traditional endpoints of survival and tumor response [9–12]. Moreover, there are relatively few examples where formal assessment of QoL, or of symptom control, has influenced clinical decision-making.

The present article provides a critical review of how QoL and/or symptom control endpoints have been used in published RCTs for patients with advanced cancer. The goal is to learn from this prior experience and to suggest ways to improve the assessment of QoL and symptom control in clinical trials where one of the principal aims is to provide palliation to patients with cancer. The importance of defining response for individual patients, rather than group changes, is highlighted.

methods

literature search

In February 2004, a literature search was performed using MEDLINE and the Cochrane Central Register of Controlled Trials (Cochrane Library 2004,
First quarter) using the following key words from the title and/or the abstract: health-related quality of life AND advanced cancer, and/or palliation, symptoms and/or metastatic cancer. The search strategy was restricted to ‘randomized controlled trials’ where the subjects were greater than 18 years old and to publications in English dating from January 1994 to January 2004.

We reviewed randomized phase III trials that assessed self-reported QoL or symptom control as a primary or secondary outcome measure in patients with advanced cancer. We were interested in larger trials and so set an arbitrary lower limit on sample size of 150 patients. Trials that used only physician-reported questionnaires were excluded, as were studies that did not evaluate a palliative strategy or treatment.

**Assessment of quality of the studies**

As there are no agreed international standards for assessing QoL or symptom control in RCTs, we designed a data abstraction form based on published guidelines that assess the quality of reporting of QoL. [11, 13, 14]. Our major criteria were selected from the Minimum Standard Checklist for evaluating health-related QoL in cancer clinical trials developed for prostate randomized clinical trials by Efficace et al. [11]. This checklist was designed to consolidate standards of reporting trials (CONSORT) including minimum standards for evaluating QoL, conceptual QoL issues (e.g. a priori hypothesis, rationale for selection of the instrument used), reporting the validity of the measurement, the methodology (instrument administration and describing timing of assessments, documenting missing data) and the interpretation of the data (clinical significance and presentation of the results) [11, 13]. In their review, properties of QoL assessment (psychometric properties, cultural validity, adequacy of domains) were documented in >80% of studies, but the reporting of an a priori hypothesis and clinical significance was rare (<15%). Furthermore, no item focused on the definition of the palliative hypothesis, the palliative response and the analysis of QoL and/or symptom response, which are major concerns in palliative clinical trials. We concentrated therefore on the choice of an a priori palliative hypothesis and other criteria such as definition of a primary QoL or symptom control endpoint, clinical significance and presentation of results.

We determined how many trials analyzed QoL results according to our pre-defined criteria. We selected 10 criteria that we believe should be regarded as a minimum checklist when evaluating the quality of a RCT and that should be reported in future RCTs that assess QoL and symptom control. The criteria are as follows:

(i) Inclusion of an a priori statistical hypothesis for improvement of QoL or symptom control.
(ii) Use of validated questionnaire(s) with established and published psychometric properties.
(iii) Designation of a specific symptom or QoL domain as the primary ‘palliative’ endpoint.
(iv) Provision of a precise definition of palliative response. (For example, a pre-defined change in a QoL score, or a pre-defined decrease in pain or other symptom).
(v) Inclusion of a plan for statistical corrections when multiple outcome measures are used, to minimize the risk of reporting a difference by chance.
(vi) Inclusion of a plan to manage missing data.
(vii) Analysis of QoL or symptom control scores for individual patients and not just a comparison of means of scores between two groups at different time points.
(viii) Statement of the proportion of patients who achieved a pre-defined palliative response.
(ix) Estimates of the duration of symptom control or palliative response. (A palliative response obtained for 1 month may not have the same significance as one of longer duration).
(x) Discussion of the limitations of the results (e.g. limitations of the tools used and statistical considerations).

Two authors (F.J. and J.V.), who were not involved in any of the identified studies, audited all the articles. The reviewers assessed the first 10 articles together, then 10 articles independently followed by comparisons of the results; disagreements were resolved by discussion. For subsequent articles any unclear issues were discussed and a consensus was reached. We examined the methods and reporting of QoL and/or symptom control in each article, based on previously published criteria and our 10-point checklist.

**Results**

Initially 182 studies were identified but this number was reduced to 112 after reviewing the articles and removing those that did not comply with the search criteria (58 were excluded because of a sample size of <150, six were the same studies published in different journals and six were not original RCT’s). Five articles required discussion between the two reviewers to resolve disagreements. The characteristics of the studies are listed in Table 1; references to them are provided as Supplementary data to the present article. Approximately half the studies were published after 2000. The median sample size was 296 patients. Sixty-three percent of the studies included a chemotherapy regimen and 18% hormonal therapy as the experimental arm.

**Analysis of methods**

Methods used in the 112 studies are summarized in Table 2. More than one QoL or symptom control questionnaire was selected.

**Table 1. Characteristics of the studies**

<table>
<thead>
<tr>
<th>Number of trials (n = 112)</th>
<th>Year of publication</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1994–9</td>
</tr>
<tr>
<td></td>
<td>2000–January 2004</td>
</tr>
<tr>
<td>Journal</td>
<td></td>
</tr>
<tr>
<td>Journal of Clinical Oncology</td>
<td>51 (46)</td>
</tr>
<tr>
<td>Journal of the National Cancer Institute</td>
<td>8 (7)</td>
</tr>
<tr>
<td>Annals of Oncology</td>
<td>7 (6)</td>
</tr>
<tr>
<td>Other</td>
<td>46 (41)</td>
</tr>
<tr>
<td>Sample size</td>
<td></td>
</tr>
<tr>
<td>150–200</td>
<td>17 (15)</td>
</tr>
<tr>
<td>201–500</td>
<td>76 (68)</td>
</tr>
<tr>
<td>&gt;500</td>
<td>19 (17)</td>
</tr>
<tr>
<td>Cancer type</td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>31 (28)</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>29 (26)</td>
</tr>
<tr>
<td>Genitourinary</td>
<td>20 (18)</td>
</tr>
<tr>
<td>Breast</td>
<td>17 (15)</td>
</tr>
<tr>
<td>Other</td>
<td>15 (13)</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>70 (63)</td>
</tr>
<tr>
<td>Hormone therapy</td>
<td>20 (18)</td>
</tr>
<tr>
<td>Supportive treatment</td>
<td>19 (17)</td>
</tr>
<tr>
<td>Other</td>
<td>12 (11)</td>
</tr>
</tbody>
</table>

*Some studies included more than one treatment.
A measure of QoL or symptom control was the primary endpoint in 25 trials (22%): symptom control in nine, QoL in 13 and a composite of the two measures in three studies. It was both a primary and secondary endpoint in 13%, and a secondary endpoint only in 78%. A hypothesis relating to improvement in QoL or symptom control was presented in only 21 studies (19%) and just 16 studies (14%) calculated sample size based on a QoL or symptom control hypothesis. An a priori QoL hypothesis was presented in 19 of the 25 studies (76%) where QoL or symptom control was a primary endpoint, and in only 13% of studies when it was a secondary endpoint. A statistical plan for the analysis of QoL/symptom control data was described in only 14% of the studies where QoL or symptom control was a secondary endpoint. A minimal change in a QoL or symptom control endpoint that would be considered clinically meaningful was defined in only 21% of the studies. Fewer than one-quarter of the trials defined a ‘palliative response’ in terms of QoL or symptom control. Very few studies (16%) considered the statistical issue of missing data and only 10% of the studies made statistical adjustments when multiple comparisons were made.

Table 4 is a summary of the studies that had classical primary endpoints (overall survival, progression-free survival or response rate). Few studies showed a significant difference in the primary endpoint. When such a difference was found it was very small: for example, the difference in median overall survival was between 0.3 and 3 months.

Table 2. QoL instruments and symptom control questionnaires

<table>
<thead>
<tr>
<th>Number of questionnaires used per study (%) (n = 112)</th>
<th>Number of studies (%) (n = 112)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>57 (51)</td>
</tr>
<tr>
<td>2</td>
<td>50 (45)</td>
</tr>
<tr>
<td>≥3</td>
<td>5 (4)</td>
</tr>
<tr>
<td>Validated QoL questionnaires used per studya</td>
<td></td>
</tr>
<tr>
<td>All questionnaires validated</td>
<td>76 (68)</td>
</tr>
<tr>
<td>Some questionnaires validated</td>
<td>19 (17)</td>
</tr>
<tr>
<td>No questionnaire validated</td>
<td>17 (15)</td>
</tr>
<tr>
<td>Types of questionnaires usedb</td>
<td></td>
</tr>
<tr>
<td>Generic questionnaires</td>
<td></td>
</tr>
<tr>
<td>EORTC-QLQ-C30</td>
<td>53 (47)</td>
</tr>
<tr>
<td>FACT-G</td>
<td>10 (9)</td>
</tr>
<tr>
<td>Rotterdam symptom checklist</td>
<td>10 (9)</td>
</tr>
<tr>
<td>Other generic questionnaires</td>
<td>25 (22)</td>
</tr>
<tr>
<td>Specific questionnaires</td>
<td></td>
</tr>
<tr>
<td>EORTC modules</td>
<td>14 (13)</td>
</tr>
<tr>
<td>FACT modules</td>
<td>9 (8)</td>
</tr>
<tr>
<td>Pain scales</td>
<td>12 (11)</td>
</tr>
<tr>
<td>Other symptom-specific modules</td>
<td>39 (35)</td>
</tr>
<tr>
<td>Description of the properties of QoL questionnaire in the article</td>
<td>104 (93)</td>
</tr>
<tr>
<td>Methods of administration reported</td>
<td></td>
</tr>
<tr>
<td>Timing of assessments</td>
<td>105 (94)</td>
</tr>
<tr>
<td>Method of administration of questionnaires</td>
<td>87 (78)</td>
</tr>
</tbody>
</table>

aQuestionnaires with acknowledged, published psychometric properties. bSome studies used more than one questionnaire.

used in 49% of them. Fifteen percent of the studies did not use any validated QoL questionnaire. A higher percentage of validated questionnaires was used after the year 2000 (97%), compared with earlier publications (72%). The majority of the studies used a generic cancer questionnaire: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire—Core 30 (EORTC-QLQ C30) [15] (47%); Functional Assessment of Cancer Therapy—General (FACT-G) [16, 17] (9%); Rotterdam symptom checklist (9%). Specific questionnaires were also used frequently, either alone or in association with generic questionnaires: 13% used EORTC modules, 8% FACT modules, 11% a variety of scales assessing pain and 35% various symptom-specific checklists. All articles, except two that evaluated symptom control, included at least one QoL questionnaire. Most articles reported the type of QoL/symptom control questionnaire used (93%), but only 26% of articles identified the questionnaire in the abstract. The timing of the assessments was well reported (94%) and 78% of authors included a description of how the questionnaire was administered. The quality of information improved with time, with 85% describing the method of administration of the questionnaires in and after the year 2000 compared with 70% before 2000.

The endpoints and statistical analyses used in the studies are summarized in Table 3. Thirty-seven percent of the studies had survival or progression-free survival as the primary endpoint.
evaluation of reported results

Results of the reviewed studies are summarized in Table 5. The majority of studies (61%) reported patient compliance for completion of the QoL or symptom control questionnaires. Comparison between the groups of patients of mean or median scores for a QoL or symptom scale was the most frequent mode of presentation of the data (81% of the studies reviewed); changes of scores for individual patients were reported less often (25%). The proportion of patients that obtained a palliative response was described in 21% of studies, the duration of response in 13% and confirmation of the response at a subsequent assessment in only 4%. Limitations of the results for QoL or symptom control were discussed in 51% of the studies.

trials meeting criteria of our checklist

The studies were analyzed in accordance with the 10 criteria listed in Methods, and the frequency of each criterion reported in the articles is summarized in Table 6. None of the studies met all 10 criteria. Only 15% met six to nine of the criteria, 26% three to five of the criteria and 59% met two or fewer. Six of the studies met none of the criteria. Studies published since the year 2000 that met five or more of the criteria were 24%, compared with 19% published before 2000 ($P = 0.53$).

discussion

Assessment of QoL and symptom control is time-consuming and expensive, and it is therefore important to obtain data that have the potential to influence clinical decision-making [18]. All of the RCTs reviewed for this article were designed for patients with advanced cancer with one of the major aims being palliation; therefore assessment of QoL and symptom control was important. Based on available guidelines [11, 13] we reviewed how QoL and symptom control were assessed in these RCTs. The criteria for our literature search (English language restriction, large sample size, limited key words) were not exhaustive and may have omitted some well-conducted QoL studies (e.g. those in other languages [18]). The objective was to ascertain how QoL and symptom control had been performed and to propose guidelines to improve the integration of these outcomes in future RCTs.

QoL or symptom control was used rarely as a primary endpoint and few studies provided high-quality information. This highlights the difficulty that many oncologists have in designing trials for patients with advanced disease, such that there is an inappropriate focus on the unrealistic expectation of obtaining a meaningful increase in survival time. It is easier to recruit patients with an incurable disease to a study focused on increasing median survival by 2 months than a study designed to improve QoL or symptoms.

There is confusion between assessment of symptom control and of QoL, which are complementary but different concepts. QoL refers to a subjective, multidimensional concept that includes psychological, social and physical domains, whereas symptom control refers to alleviation of one or more symptoms. Assessment of symptoms requires self-report symptom scales or checklists: the most commonly used are pain scales, which are often associated with an objective measure of analgesic consumption. Evaluation of QoL is performed with questionnaires that include different domains of QoL (such as physical, psychological and social domains). Clinical studies that evaluate the impact of treatments on symptom control usually add a QoL assessment.

As discussed below, most of the reviewed articles used suboptimal methodology. For this reason we selected 10 criteria that might be incorporated into clinical trials to improve the role of QoL and/or symptom control endpoints in making...
Table 6. Proposed checklist of 10 important criteria to be reported in RCTs assessing QoL and/or symptom control as major endpoints in advanced cancer, and the number of trials that included these criteria

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Percentage of trials (n = 112)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective of study</td>
<td></td>
</tr>
<tr>
<td>1. A priori hypothesis relating to QoL or symptom control Methodology</td>
<td>19</td>
</tr>
<tr>
<td>2. Use of a validated questionnaire</td>
<td>85</td>
</tr>
<tr>
<td>3. Definition of a primary palliative endpoint</td>
<td>22</td>
</tr>
<tr>
<td>4. Definition of a palliative response</td>
<td>24</td>
</tr>
<tr>
<td>5. Statistical corrections when multiple comparisons made</td>
<td>10</td>
</tr>
<tr>
<td>6. Missing data described and analyzed</td>
<td>16</td>
</tr>
<tr>
<td>Results and discussion</td>
<td></td>
</tr>
<tr>
<td>7. Comparison of changes in QoL or a symptom scale for individual patients (instead of comparison of mean scores for QoL)</td>
<td>25</td>
</tr>
<tr>
<td>8. Proportion of patients who achieved a palliative response</td>
<td>21</td>
</tr>
<tr>
<td>9. Duration of palliative response</td>
<td>13</td>
</tr>
<tr>
<td>10. Discussion of limits of QoL results</td>
<td>51</td>
</tr>
</tbody>
</table>

The choice of questionnaire

It is accepted that evaluation of QoL or symptom control should be based on patient self-report [19, 20]. Physicians are poor judges of patients’ QoL; many studies have shown that there are important discrepancies between toxicities reported according to the Common Toxicity Criteria by physicians, and patients’ self-reported toxicity ratings [20, 21]. In our review, 85% of the studies used at least one validated self-report questionnaire with an improvement in such use over time, and >90% described the questionnaire used, in agreement with the results of Efficace et al. [10]. There has been improvement in the use of QoL instruments in RCTs, and clinicians now have greater knowledge of the different instruments. The questionnaires were identified in the abstracts of only 26% of the articles, and we suggest that questionnaires defining a major palliative endpoint should be identified in the abstracts of future papers. Forty-nine percent of the studies used more than one questionnaire, most used a generic questionnaire [15, 16], and 35% used a disease-specific module (e.g. breast, lung, prostate, etc.) or symptom scale (e.g. fatigue and pain). Inclusion of both generic and specific questionnaires for QoL and symptom control is particularly appropriate for most clinical trials that seek to evaluate palliation, as symptom modules often add information that is not provided by generic questionnaires [22].

QoL/symptom control endpoints

In the studies we reviewed one of the major aims was palliation, yet a minority (22%) used QoL or symptom control as a primary endpoint. Where symptom control was the main endpoint, QoL results were generally (and appropriately) reported as descriptive data, with changes in scores for the different domains over time. A composite primary endpoint (including symptom assessment and QoL) might also be appropriate, although minimal data are available on the relationship between health-related QoL and symptom control. Careful attention needs to be given to the selection of palliative outcome measures for clinical trials [23].

In most of the studies, a measure of QoL or symptom control is used as a secondary endpoint. Only 19% of the studies stated a hypothesis related to QoL or symptom control, and only 21% defined the size of the difference that would be considered clinically significant. A sample size based on a primary QoL/symptom control hypothesis can be calculated as for other endpoints such as overall survival, and can incorporate realistic adjustments for death, non-compliance or drop-out due to patients being too unwell to participate [24, 25]. In most of the reviewed studies sample size was calculated based on the primary endpoint; many of them may not therefore have been adequately powered to determine clinically important changes in QoL or symptom control. Of 33 with overall survival as their primary endpoint, 13 studies demonstrated a survival advantage (ranging from a few weeks to 3 months), and only four demonstrated an improvement in QoL. One possible reason may be the lack of power to show a difference in QoL. Palliation should imply an improvement in either duration and/or quality of survival, and it is desirable to use composite or co-primary endpoints that include survival and an a priori, well-defined QoL and/or symptom control.

Most of the studies presented multiple comparisons, yet only 10% of articles included statistical adjustment for them. The risk of finding a difference when a true difference does not exist (i.e. type I error) is increased by statistical comparison of multiple domains or item scores of the multidimensional instruments [26]. Analyses of planned comparisons and statistical correction should be defined initially.

Few studies defined a strategy for handling missing data, an important issue in studies of patients with advanced cancer. Data are generally not missing at random and therefore bias can be introduced [27–30]. The benefit of a palliative intervention may be overestimated by comparison of group means as only individuals who remain well enough to participate in questionnaires provide data. Missing data become less of a problem if palliative response is defined by serial measurements for individual patients, as described below, since non-returning patients are then regarded as not satisfying the criteria for a palliative response.

Presentation of QoL/symptom control data: the importance of defining response for individual patients

In our review, changes from baseline in QoL or symptom control scores were reported for individual patients in only...
25% of the studies. More than 80% of the studies reported changes in group means or medians, with an average determined at baseline for each group, and then at one or more pre-defined times of follow-up. This strategy is sub-optimal for assessment of QoL or symptoms in clinical trials [29, 31]. QoL is the property of an individual and not of a group; it will improve in some patients and decline in others, and will do so at variable rates, so that average values may have little meaning [32]. A strategy that leads to marked improvement in QoL in 50% of patients with an advanced malignancy, while QoL declines in the other 50% due to progression of disease, would represent substantial therapeutic effect, but a comparison of mean QoL scores with baseline might show no difference. Also, some patients will drop out (usually those who are doing poorly), and comparison of a reduced sample at varying times on treatment with the whole sample at baseline, or with the averaged score at baseline for the patients that remain on study, leads to bias that confounds interpretation.

The above problems can be overcome by defining criteria for ‘palliative response’ for individual patients and determining the proportion of patients that satisfy it [2, 7, 33]. Analogy with tumor response can point to a better way for analyzing symptoms or QoL. To use QoL or a symptom scale as a major outcome, we need to define a simple, relevant primary measure of QoL or symptom control, and to establish a hypothesis (preferably supported by data) about the magnitude of change in this endpoint that will be considered clinically meaningful [30, 31, 34–36]. One can also define when this minimal improvement in the primary measure has occurred, and its duration, analogous to definition of tumor response and its duration by RECIST or other criteria [37]. Osoba et al. [30] concluded that a change of 5–10 points in a normalized QoL scale represented a small clinical change, a change of 10–20 points represented a moderate change, and a change of >20 points represented a large change. Thus, setting palliative response as a 10- or 15-point change in a normalized QoL or symptom scale seems reasonable. The QoL or symptom

<table>
<thead>
<tr>
<th>SUGGESTED APPROACH</th>
<th>STRATEGIES TO AVOID</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition of Palliative Endpoints</strong></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td></td>
</tr>
<tr>
<td>• Improvement by ≥15 points on a normalized scale of a pain score, without increase in anaglesic consumption, confirmed on 2 consecutive occasions at least 3 weeks apart.</td>
<td>• Lack of pre-defined primary and secondary endpoints for QoL or symptom control</td>
</tr>
<tr>
<td>• Overall survival</td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
</tr>
<tr>
<td>• Improvement of ≥ 15 points on a normalized global score of a self-report, cancer specific validated QoL scale (e.g. FACT-G).</td>
<td></td>
</tr>
<tr>
<td>• Improvement of ≥ 15 points on a normalized fatigue scale (e.g. FACT-F).</td>
<td></td>
</tr>
<tr>
<td><strong>A Priori Palliative Hypotheses</strong></td>
<td></td>
</tr>
<tr>
<td>• Treatment A will lead to ≥15% more patients satisfying the primary criterion of palliative response than Treatment B.</td>
<td>• Mean QoL will be better for treatment A compared to treatment B at 3 months.</td>
</tr>
<tr>
<td>• Median survival will be 3 months longer for patients receiving Treatment A</td>
<td></td>
</tr>
<tr>
<td><strong>Sample Size</strong></td>
<td></td>
</tr>
<tr>
<td>• Sample size calculated on the basis of the &quot;palliative&quot; hypothesis (and adjusted for the second primary endpoint of overall survival).</td>
<td>• Sample size based only on endpoint for survival or tumor response</td>
</tr>
<tr>
<td><strong>Presentation of the Results</strong></td>
<td></td>
</tr>
<tr>
<td>• Proportion of patients who achieve a palliative response is compared between group A and B</td>
<td>• Mean scores of QoL are compared between the 2 groups</td>
</tr>
<tr>
<td>• Distribution of duration of palliative response is reported for each arm</td>
<td>• No duration of response is given</td>
</tr>
<tr>
<td>• Actuarial survival curves</td>
<td></td>
</tr>
<tr>
<td><strong>Other Symptom Control or QoL Data</strong></td>
<td></td>
</tr>
<tr>
<td>• Statistical correction planned for comparison of scores for secondary endpoints between the 2 groups (e.g. Bonferroni correction)</td>
<td>• Comparison of mean values for multiple items in QoL or symptom scales between groups A and B without correction</td>
</tr>
<tr>
<td>• Changes in others dimensions of QoL questionnaire reported for individual patients as descriptive data</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. An illustration of a hypothetical RCT designed to compare palliative chemotherapy regimens for a common type of metastatic cancer where pain and fatigue are the major symptoms.
endpoint is then assessed repeatedly with time (like tumor size) and the proportion of patients who satisfy ‘palliative response’ is defined by those who satisfy the pre-defined criterion for a minimum period of time—typically on two assessments at least 3–4 weeks apart. As for tumor response, one can also define the median and distribution for duration of ‘palliative response’, which will vary among the patients, and define criteria for progression in the QoL or symptom scale endpoint. This method is not influenced by drop-out (again similar to tumor response) since an intention-to-treat analysis is used in which those patients who do not return for assessment are classified as non-responders [35, 38]. We found very few studies that followed this method: 21% reported a definition of palliative response, only 13% reported the duration of the palliative response and 4% confirmed that a palliative response had occurred by repeat assessment 3–4 weeks later.

In Figure 1 we provide an illustration of our proposed strategy for analysis of QoL and symptom control for a hypothetical RCT, as well as suggesting strategies to avoid.

conclusion

Although there has been improvement during the last decade in the inclusion and assessment of QoL and symptom control in RCTs, these data have contributed but rarely to more classical endpoints in decision-making, mainly because of sub-optimal methodological standards. Greater effort is required to establish outcome criteria and guidelines for reporting data in RCTs that focus on QoL and symptom control in patients with advanced cancer. These criteria should be applied with the same rigor that is applied to traditional endpoints. Application of our 10-point checklist might help to improve the standard of reporting of data for QoL and symptom control.

acknowledgements

We wish to thank Cindy Ho for performing data entry for this study.

references


