ELSEVIER

European Journal of Cancer 41 (2005) 280–287

European Journal of Cancer

www.ejconline.com

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

David Osoba [a,*], Andrea Bezjak [b], Michael Brundage [c], Benny Zee [d], Dongsheng Tu [e], Joseph Pater [e], for the Quality of Life Committee of the NCIC CTG

[a] *QOL Consulting, 4939 Edendale Court, West Vancouver, Canada BC V7W 3H7*
[b] *Radiation Oncology, Princess Margaret Hospital, University of Toronto, Toronto, Ont., Canada*
[c] *Radiation Oncology, Queen's University, Kingston, Ont., Canada*
[d] *Chinese University of Hong Kong, Shatin, Hong Kong, China; Former Senior Statistician, NCIC CTG, Kingston, Ont., Canada*
[e] *NCIC CTG, Queen's University, Kingston, Ont., Canada*

Received 8 June 2004; received in revised form 18 October 2004; accepted 19 October 2004
Available online 19 November 2004

Abstract

Clinicians are being confronted with increasing amounts of health-related quality of life (HRQOL) data. Thus, there is a need for a greater understanding of the analysis and interpretation of HRQOL so that the data can be reported in an appropriate manner. The approach of the Clinical Trials Group of the National Cancer Institute of Canada emphasises the clinical meaning of the results, while avoiding complex statistical modelling. It consists of four steps: calculating the questionnaire completion rates, calculating the baseline scores, determining the individual change in scores over time for the domains specified in the trial hypothesis, and culminates in determining the proportions of patients who have reported clinically meaningful changes in scores since baseline. A rationale supporting each step is given. This approach is presented as a simple and practical aid to the analysis, interpretation and reporting of HRQOL results.

Keywords: Health-related quality of life; Analysis; Interpretation; Clinical trials

1. Introduction

Many publications have addressed methods for collection, analysis and interpretation of health-related quality of life (HRQOL) data [1–9]. While these are useful to researchers and clinical trials investigators devoted to HRQOL research, some methods are published in journals or books that are not easily accessed by clinicians, or present statistical methods that are difficult for physicians to use.

There have been criticisms of the adequacy of the reporting of published HRQOL results, but some of these may be premature in that they are based primarily on older studies [10–12]. Recent publications seem to show an improvement in the quality of the data and its analysis [13,14]. Nevertheless, the criticism has served to highlight the need to collect, analyse and interpret HRQOL data in a robust and scientifically sound manner, and to present the results in ways that are clinically meaningful.

These observations indicate that it is appropriate, at this time, to suggest a simple and succinct, scientifically sound and clinically oriented approach for reporting HRQOL findings. The Clinical Trials Group (CTG) of

* Corresponding author. Tel.: +1 604 921 7793; fax: +1 604 921 7794.
E-mail address: david_osoba@telus.net (D. Osoba).

0959-8049/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2004.10.017

the National Cancer Institute of Canada (NCIC) has developed an analysis approach, detailed below, during 17 years of experience in conducting HRQOL research. The approach consists of four steps that culminate in a determination of the proportions of patients that achieve a response in prespecified HRQOL domains. It is a simple and useful framework for analysis and interpretation that is intended for clinicians who wish to understand better HRQOL results.

2. Preparing for the analysis of HRQOL data

The quality of the results of the analysis of HRQOL assessments can be only as good as the quality of the data that is collected. A number of authors have suggested strategies to ensure the quality of collected HRQOL data, and we refer the reader to these discussions and reviews [1,2,6,8,9,15,16].

3. Patients and methods

3.1. The analysis steps

As a preliminary step in the analysis, the scoring procedures provided with the questionnaire being used should be followed when preparing the data. In addition, a description of the demographic characteristics of the patients in the trial should be provided and the balance of these characteristics between treatment arms ascertained.

The following steps are intended as a general guideline and may not apply to all situations. The first three steps are intended to provide the data that is required to carry out the fourth step, which is the considered to be the ultimate goal of the analysis. The steps are first described briefly below and then the rationale for each step is discussed in the next section.

3.2. Step 1 – calculating completion rates

Calculate completion rates as follows for each treatment assignment group:

(a) number of patients completing the baseline (pre-treatment) assessment over the total number of eligible patients entered,
(b) number of patients completing the assessments at designated time points (windows) while on the study over the total number entered (the ''intent-to-treat'' population),
(c) number of patients completing assessments at designated time points over the total number completing the assessment at baseline (the ''efficacy'' population),
(d) number of patients completing assessments at designated time points over the total number of patients still on study and expected to complete at each time point (the ''number expected'' population),
(e) proportion of questionnaires completed outside the designated time windows,
(f) proportion of domains and items completed within questionnaires at baseline and designated time points.

3.3. Step 2 – comparing baseline scores between groups

Calculate the mean or median baseline scores for each of the HRQOL components (domains and single items) within the questionnaire for each of the treatment groups, as follows:

(a) number of patients providing responses,
(b) mean (or median, if more appropriate) score and standard deviation (or inter-quartile range) for each HRQOL component,
(c) determine if there is an apparent difference between the mean or median scores between the treatment groups according to preset criteria.

3.4. Step 3 – comparing the change scores between- and within-treatment groups

Determine the change-from-baseline scores within the chosen time window at each specific time for each HRQOL component of interest that was specified in the hypothesis, as follows:

(a) subtract the baseline score for each individual from his/her score at each designated time point while on study,
(b) calculate the means for the differences (the mean change score) ± the standard error (SE) at each designated time point,
(c) test for statistically significant differences in the mean change scores between treatment groups according to preset criteria, and
(d) test for statistically significant differences between baseline scores and scores from subsequent designated time points within each treatment group.

3.5. Step 4 – determining the proportions of patients with improved, stable and worsened scores

Decide, *a priori*, the magnitude of change (cut-off point) that will be considered to be a clinically meaningful change in HRQOL scores and the duration that this

change will need to persist in order to consider the HRQOL response as being "improved", "worsened" and "stable". Calculate the proportions of patients with clinically meaningful change, as follows:

(a) determine the numbers of patients who reported the preset magnitude of change in scores ("improved", "worsened", or neither, i.e., "stable" response) for each domain during the study period in each treatment arm,
(b) calculate the proportions of patients who report improvement or worsening in each intent-to-treat group,
(c) calculate the median (or mean) duration of improvement in the improved group and, similarly, the duration of stable HRQOL status in the stable group, and
(d) test for statistically significant differences between the three categories of responses.

4. Results

4.1. Rationale and explanatory notes for the analysis steps

4.1.1. Step 1 – calculating completion rates

Collecting HRQOL assessments at baseline, i.e., just before the beginning of an intervention or treatment is important for a determination of the balance between treatment groups at baseline. Any significant imbalance must be adjusted for in the subsequent analysis. In longitudinal studies, when HRQOL will be measured at multiple time points, the baseline assessment scores will serve as a comparison for within-group changes in HRQOL scores over time. In addition, baseline scores can be used for between-trial and between-disease comparisons. Baseline assessments should be carried out as close to the beginning of the intervention as is reasonably possible to minimise any possible changes in HRQOL scores between the assessment and the intervention.

(a) Ideally, the number of patients completing the baseline assessment should be identical to the number of patients enrolled in the trial. Missing data at baseline will result in a loss of information about the possible efficacy of therapy, even if all the subsequent completions are available. The reasons for failure to collect all the baseline data should be specified in detail.
(b) The rates of completed questionnaires at specified "time windows" allows a calculation of the proportions of enrolled patients for whom data exist at time points along the trial and will show the rate of dropout from the study over time. These numbers will also serve as the numerator for (d), below.
(c) The proportions of patients with baseline data that complete on-study assessments will be the same as those in (b) above, if all patients completed HRQOL assessments at baseline. When the number completing at baseline is less than the number enrolled in the trial, the interpretation of the HRQOL analysis is subject to limited generalisation beyond the participating patients.
(d) The proportions of patients who complete questionnaires over the number expected to complete the questionnaires at each designated time point will distinguish between failure to obtain completed questionnaires despite patients still remaining on study, and failure because patients are no longer on study due to disease progression or death. Low completion rates and dissimilar proportions in each treatment arm will jeopardise the interpretation of the study and its generalisability. Fig. 1 illustrates the overall completion rates over time with

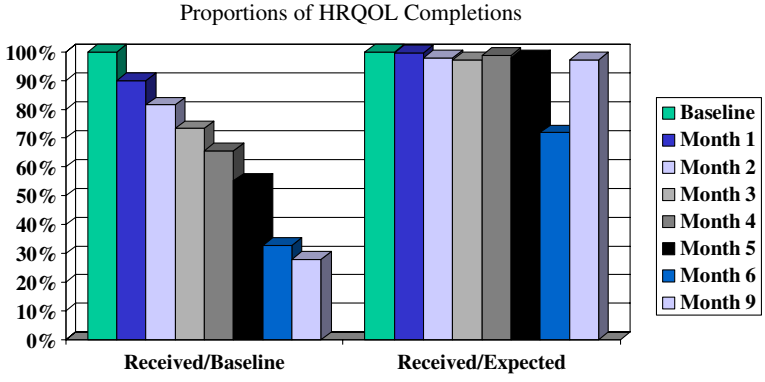

Fig. 1. Proportions of health-related quality of life (HRQOL) questionnaires completed at baseline and during the study. The received/baseline proportions show a marked attrition of questionnaires over time with only 27% completion by month 9. However, the received/expected proportions remained high at >95% for all but month 6 of the study. (Adapted from Sadura A and colleagues in [17].)

respect to the number of baseline completions (left side of graph) and the number expected to be completed (right side of graph) in a randomised clinical trial of two forms of chemotherapy [17]. The number of completions received over the number of baseline completions drops off rapidly over time, but the patients still alive during the study are completing questionnaires at a high rate. Thus, there is very little randomly missing data.

(e) Reporting the number of assessments that did not fall within specified time windows may explain why there are fewer than the expected number of assessments that are subjected to analysis. These assessments should not be analysed together with those that fall within the windows, except in growth curve analyses [7].

(f) The proportion of competed domains and items within questionnaires should be noted for each of the components above, except for (e).

Completion of all the components of Step 1 allows the monitoring of completion rates at the level of the institution, the patient, and the questionnaire while on treatment and during follow-up after treatment [7,17]. This complete data-set has rarely been presented in reports from clinical trials [18]. Yet, it is important for the critical interpretation of the reported results of HRQOL analyses.

4.1.2. Step 2 – comparing baseline scores between groups

The calculations in this step are straightforward and provide the foundation to which scores at subsequent time points will be compared. Mean or median baseline scores for the domains/items of interest that vary significantly between treatment groups suggest that they are dissimilar and will lead to potential difficulties in the interpretation of the results.

4.1.3. Step 3 – comparing the change scores between- and within-treatment groups

The comparison of the mean scores for each treatment group are commonly reported and are appropriate when there is almost no attrition in patients on study. However, when there has been significant attrition in the data over time, comparing group mean scores at each time point with the group mean scores of all the patients at baseline can lead to spurious results because of a "survivor effect". Fig. 2 is a comparison of group mean scores over time in a clinical trial [19]. Less than half of the patients completing questionnaires at baseline were still available at time 3. Global QOL scores improved over time in both treatment groups, but it is unknown whether this was the result of a true treatment effect or of a survivor effect.

One solution is to limit the analysis of the HRQOL changes to those patients who provide data both at baseline and at the subsequent time point of interest. Doing so will eliminate spurious increases in group mean scores due to bias in the subset of patients who are able to complete data at both times (for example, if patients with lower scores are more likely to drop out). However, the solution is not ideal, since the limitation of the analysis of mean scores to this patient subset raises issues of the generalisability of results to the treatment group overall. Our preference is to calculate change scores for individual patients and to address missing data as we discuss below.

The main purpose of Step 3 is to provide change scores for individual patients that are necessary for the final analysis in Step 4. However, it is worth reporting the change scores so that an estimate of the magnitude of change in each domain can be made (for example, see Fig. 3) [20]. The magnitude of change can be tested for statistical significance.

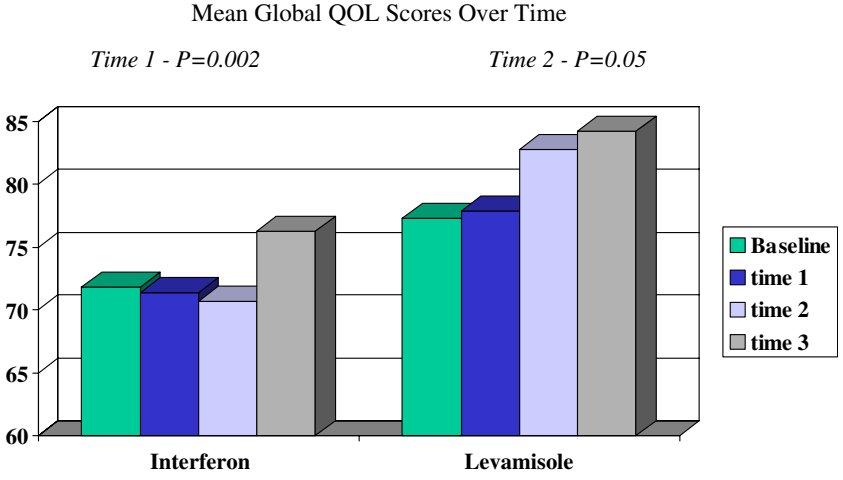

Fig. 2. Mean global quality of life (QOL) scores increase over time, but baseline scores for the patients treated with levamisole were higher than they were for the interferon group. The data illustrate a possible survivor effect for both treatment groups. (Adapted from Osoba D and colleagues in [19].)

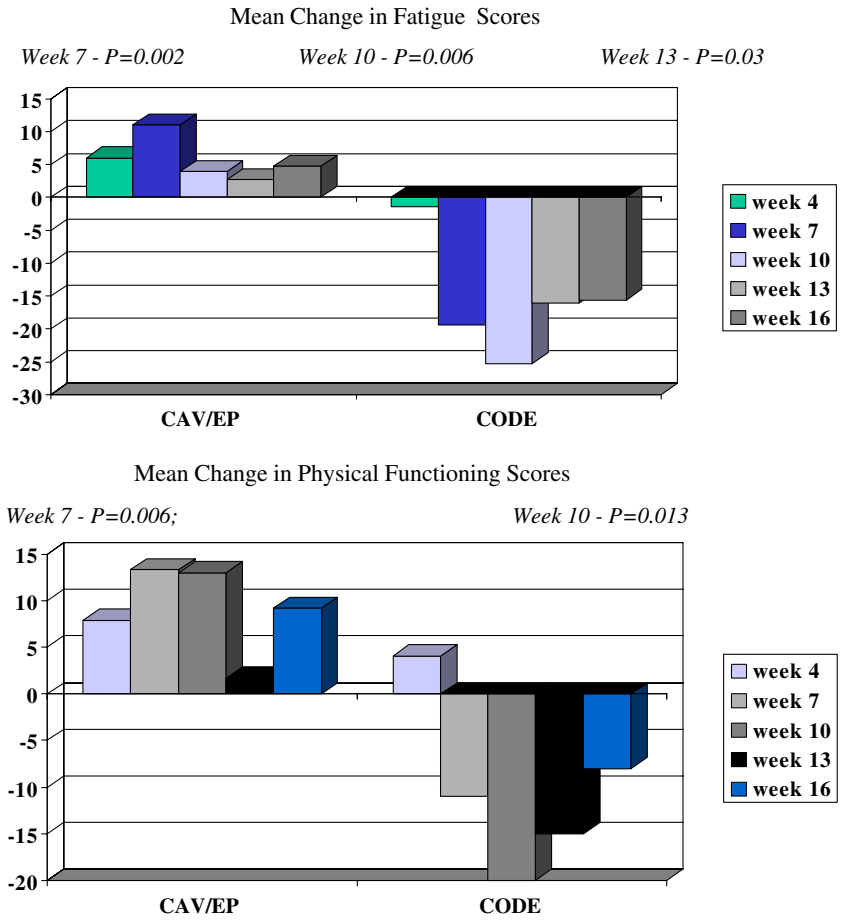

Fig. 3. Mean change in fatigue and physical functioning scores from baseline showing statistically significantly different scores between the cyclophosphamide, doxorubicin, vincristine/etoposide, cisplatin (CAV/EP) and cyclophosphamide, vincristine, cisplatin, etoposide (CODE) arms of a study in patients with small-cell lung cancer. (Adapted from Murray N and colleagues in [20].)

4.1.4. Step 4 – determining the proportions of patients with improved, stable and worsened scores

The determination of how much change constitutes a clinically meaningful difference in HRQOL scores is the subject of intense study. Some have relied primarily on statistical procedures, e.g., testing for the null hypothesis and effect size. Others are based on the properties of the measurement tool, e.g., standard error of the measurement, or the use of either external anchors or global ratings of change. Details of these approaches can be found elsewhere [5,21–25].

It is interesting to note that these different approaches, using different HRQOL measures in a variety of diseases (including several kinds of cancer), have given rise to similar answers. A change of from 5% to 10% (or in general, 0.5 of a standard deviation) of the scale breadth is perceptible to patients as a meaningful change [26–29]. It has also been designated as a clinically meaningful change by a panel of physicians with respect to two questionnaires, The Chronic Respiratory Disease Questionnaire and the Medical Outcomes Study Short Form 36 [30].

Some studies have used cut-off points as low as 5% [31]. However, 10% of the scale breadth appears to be a more reasonable number to use as a cut-off point when classifying patients into "improved", "stable" and "worsened" HRQOL categories. The 10% cut-off point is less likely to include scores that are "false positives" than is a lower cut-off point. Nevertheless, a justification should be given for the chosen cut-off point(s), whether the cut-off point is the same for "improved" and "worsened" scores, and how long the change must last. Clinicians should attempt to give examples of the clinical implications of the chosen cut-off point since this would be helpful in understanding the meaning of the change, e.g., how much more are patients whose physical functioning improved by 10% able to do, or what does a 10% change in pain scores mean to the amount of analgesia that patients require.

It should be noted that in some questionnaires, the nature of the construction and scoring of some domains/symptoms is not readily amenable to using a 10% cut-off point for calculating the proportions of patients with improvement or worsening. For example, if a

domain consists of only two items or a symptom of only one item, both with a four-category response option, then a change of one category will equal either a 17% or 33% change, respectively. This magnitude is well above the cut-off point value of 10% and may represent an underestimate of the proportions of patients experiencing the change. Methods for dealing with this situation have been suggested in [31].

(a) To be classified as "improved", the scores must improve above baseline by the cut-off point value for a specified period of time or a specified number of assessments during the study [32]. Scores that decline below baseline by the cut-off point value and never recover are classified as "worsened". Scores that are neither "improved" nor "worsened" are classified as "stable" HRQOL [21].
(b) Using the above criteria, the calculation of the number of patients in each category is straightforward.
(c) Assessing the duration of improvement provides additional informative data on the relative efficacy of the treatments being studied [32]. The end of the improvement period is reached when the scores fall to, or below, baseline values after having satisfied the preset criteria for improvement. A similar calculation assessing time to deterioration could include all patients whether or not they improved at some point in the trial.
(d) The proportions of patients in each category can be compared statistically for differences between the treatment arms using standard statistical methods. The result of a clinical trial illustrating the proportions of patients with varying proportions of responses is presented in Fig. 4. Although the group receiving treatment A consistently showed a larger proportion of patients with a $\geqslant 10\%$ improvement in each domain and symptom, only the fatigue and domains were improved statistically significantly more frequently ($P < 0.05$ after a Bonferroni correction). There was no difference in the proportions of patients showing worsened scores in the two treatment groups [21].

5. Discussion

Our approach to the analysis of HRQOL data from phase III clinical trials is relatively simple and clinically intuitive, and, therefore, appealing. Our preferred analysis consists of a step-by-step approach to determining the quality and completeness of the data, and calculating the proportion of patients with clinically meaningful differences (i.e., the proportions with improved, stable, and deteriorated scores) at specified time points in the trajectory of their treatment or follow-up.

In the analysis of group mean scores, we have preferred not to impute missing data but, at times, have used growth curve analysis because it incorporates all the data available for patients over the desired time intervals [7,20]. Others have assessed the use of imputation methods [33]. One study of several statistical methods suggests that an approach based on the linear mixed model may be a powerful method, but the assumptions underlying it need further testing before it can be fully accepted [34]. A simple approach is to ignore missing data if the proportion is small. It is uncertain if this approach is inferior to growth curves and imputation methods, and these methods need to be tested by direct comparison using data from the same trial. In the meantime, we suggest that a combination of distribution-based and anchor-based methods may be the preferred approach to analysis of data from clinical trials [5,6,8,22,25,26].

One of the main areas of continuing study has been how to deal with missing data, whether missing at random or not [4]. The choice of appropriate statistical modelling methods and imputation methods for dealing with missing data are beyond the scope of this paper. They are dealt with in detail in standard statistical textbooks [16], as well as in recent books and articles [4,6,8,33]. One advantage of our approach of calculating the "response" proportions is that one can include the proportion of patients for whom no data is available, thus providing an "intent-to-treat" approach by including all randomised patients.

Within-group comparisons over time are carried out less frequently than are comparisons between treatment groups. The value of within-group comparisons during treatment is to determine if the treatment has been associated with improvement (or deterioration) of HRQOL scores compared with the pre-treatment status. This allows the assessment of whether a treatment appears to be of value on its own. It is a valuable approach, particularly in phase II non-randomised trials [35], but also in phase III studies [21,36].

While we believe that our approach is relatively simple and useful, new approaches will undoubtedly be necessary as the measurement of HRQOL evolves. For example, the use of item–response-theory-based assessment [37,38] will probably require a reassessment of what constitutes clinically or subjectively meaningful change in clinical trials. Other methods may be more appropriate for studying population health (e.g., utility-based measures) or changes in the health of individuals in day-to-day clinical practice.

A weakness of our approach is that it may not satisfy some of the premises that underlie complex statistical modelling approaches, e.g., the effects of non-random dropout. However, a strength of our approach i.e., the reporting of the proportion of patients who report HRQOL benefit, is that it provides an answer that is

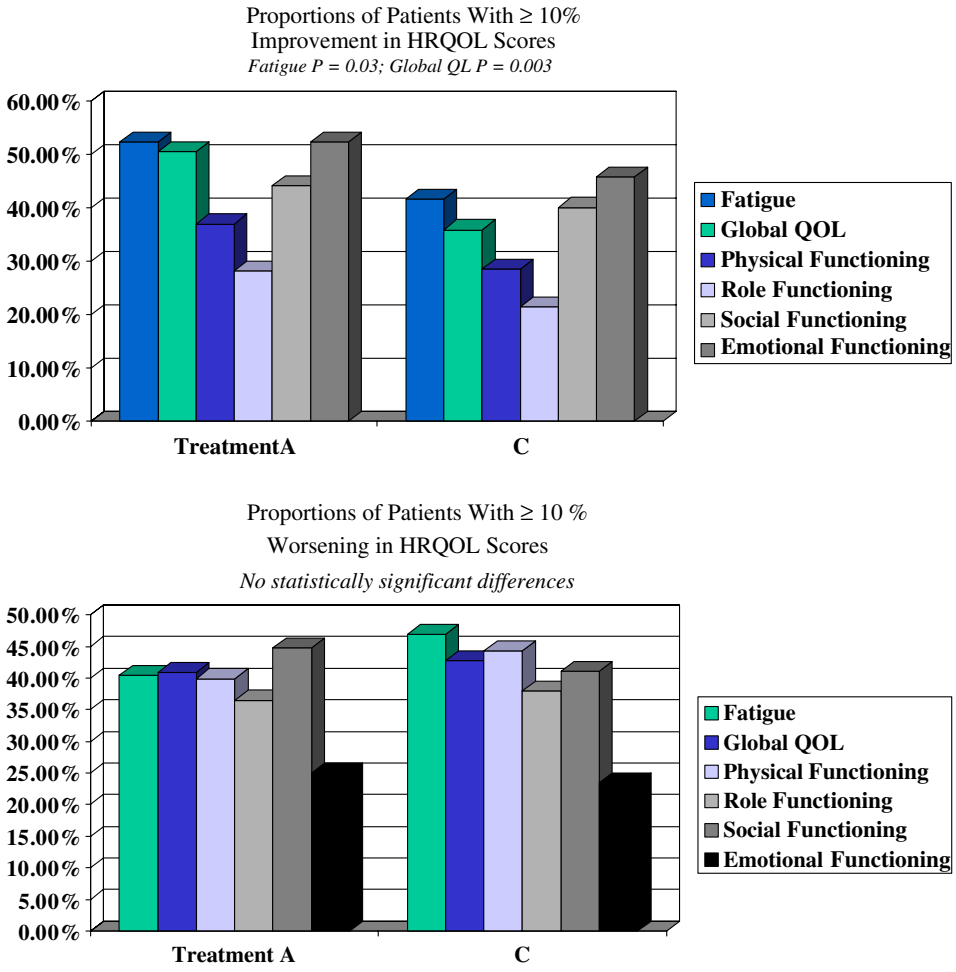

Fig. 4. Proportions of patients with breast cancer showing either improvement or worsening of HRQOL scores. The upper panel shows statistically significantly different proportions of patients with improved fatigue and global QOL scores in treatment A *vs* treatment C. (Adapted from Osoba D and colleagues in [21].)

readily meaningful to clinicians and, as a result, is more likely to influence clinical decision-making. Nevertheless, it should be stressed that HRQOL data should not be used as the sole basis for clinical decision-making. As is generally true in most of clinical medicine, a number of tools (e.g., the history, the physical examination, and various diagnostic and laboratory tests) are all integrated to derive the most likely diagnosis and to come to a decision about optimal therapy and follow-up after treatment. HRQOL data should be viewed as another useful tool in this armamentarium.

Conflict of interest statement

None declared.

References

1. Staquet M, Berzon R, Osoba D, *et al*. Guidelines for reporting results of quality of life assessment in clinical trials. *Qual Life Res* 1996, **5**, 496–502.
2. Osoba D. Guidelines for measuring health-related quality of life in clinical trials. Methods and Practice. In Staquet MJ, Hays RD, Fayers PM, eds. *Quality of Life Assessment in Clinical Trials*. Oxford, Oxford University Press, 1998. pp. 19–35.
3. 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**, 603–612.
4. Bernhard J, Gelber RD, eds. Workshop on missing data in quality of life research in cancer clinical trials: practical and methodological issues. *Stat Med* 1998, **17**, 511–796.
5. Osoba D, Rodrigues G, Myles J, *et al*. Interpreting the significance of health-related quality-of-life scores. *J Clin Oncol* 1998, **16**, 139–144.
6. Fayers P, Machin D. *Quality of Life: Assessment, Analysis and Interpretation*. Chichester, UK, Wiley, 2000.
7. Zee B, Osoba D. Health-related Quality of Life Outcomes. In Crowley J, ed. *Handbook of Statistics in Clinical Oncology*. New York, Marcel-Dekker, 2001. pp. 249–267.
8. Fairclough DL. *Design and Analysis of Quality of Life in Clinical Trials*. Boca Raton, FL, Chapman & Hall/CRC, 2002.
9. Sprangers MAG, Moinpour C, Moynihan TJ, *et al*. Assessing meaningful change over time in quality of life: a users' guide for clinicians. *Mayo Clin Proc* 2002, **77**, 561–571.
10. Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994, **272**, 619–626.

11. Leplege A, Hunt S. The problem of quality of life in medicine. *JAMA* 1997, **278**, 47–50.
12. Bottomley A, Therasse P. Quality of life in patients undergoing systemic therapy for advanced breast cancer. *Lancet Oncol* 2000, **3**, 620–628.
13. Efficace F, Bottomley A, Osoba D, *et al*. Beyond the development of health-related quality of life (HRQOL) measures: a checklist for evaluating HRQOL outcomes in cancer clinical trials – does HRQOL evaluation in prostate cancer research inform clinical decision-making?. *J Clin Oncol* 2003, **21**, 3502–3511.
14. Bottomley A, Efficace G, Vanvoorden V. What is the added clinical value of health related quality of life (HRQOL) studies in randomised controlled trials (RCTS)?. *Qual Life Res* 2003, **12**, 766., [Abstract].
15. Hays RD, Anderson R, Revicki D. Psychometric considerations in evaluating health-related quality of life measures. *Qual Life Res* 1993, **2**, 441–449.
16. Nunnally JC, Bernstein IH. *Psychometric Theory*. 3rd ed. New York, McGraw-Hill, 1994., p. 264–65.
17. Sadura A, Pater J, Osoba D, *et al*. Quality-of-life assessment: patient compliance with questionnaire completion. *J Natl Cancer Inst* 1992, **84**, 1023–1026.
18. Lee CW, Chi KN. The standard of reporting of health-related quality of life in clinical cancer trials. *J Clin Epidemiol* 2000, **53**, 451–458.
19. 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, Lippincott, 1993. pp. 412–416.
20. 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**, 2300–2308.
21. Osoba D, Slamon DJ, Burchmore M, *et al*. Effects on quality of life of combined trastuzumab and chemotherapy in women with metastatic breast cancer. *J Clin Oncol* 2002, **20**, 3106–3113.
22. Lydick E, Epstein RS. Interpretation of quality of life changes. *Qual Life Res* 1993, **2**, 221–226.
23. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989, **27**(Suppl. 3), 178–189.
24. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989, **10**, 407–415.
25. Wyrwich K, Tierney W, Wolinsky F. Further evidence supporting a SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 1999, **52**, 861–873.
26. Guyatt G, Osoba D, Wu A, *et al*. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002, **77**, 371–383.
27. Cella D, Hahn EA, Dineen K. Meaningful change in cancer-specific quality of life scores: differences between improvement and worsening. *Qual Life Res* 2002, **11**, 207–221.
28. King MT. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. *Qual Life Res* 1996, **5**, 555–567.
29. 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**, 582–592.
30. Wyrwich KW, Fihn SD, Tierney WM, *et al*. Clinically important changes in health-related quality of life for patients with chronic obstructive pulmonary disease. *J Gen Intern Med* 2003, **18**, 196–203.
31. Langendijk JA, Aaronson NK, de Jong JMA, *et al*. Prospective study on quality of life before and after radical radiotherapy in non-small-cell lung cancer. *J Clin Oncol* 2001, **19**, 2123–2133.
32. Osoba D, Tannock IF, Ernst DS, *et al*. Health-related quality of life in men with metastatic prostate cancer treated with prednisone alone or mitoxantrone and prednisone. *J Clin Oncol* 1999, **17**, 1654–1663.
33. Fairclough DL, Gagnon DD, Zagari MJ, *et al*. Evaluation of quality of life in a clinical trial with nonrandom dropout: The effect of epoetin in anemic cancer patients. *Qual Life Res* 2003, **12**, 1013–1027.
34. 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. MSc Thesis, Community Health and Epidemiology, Kingston, Ont., Queen's University, 2002.
35. Osoba D, Brada M, Yung WKA, *et al*. Health-related quality of life in patients with anaplastic astrocytoma during treatment with temozolomide. *Eur J Cancer* 2000, **36**, 1788–1795.
36. Osoba D, Brada M, Yung WKA, *et al*. Health-related quality of life in patients treated with temozolomide versus procarbazine for recurrent glioblastoma multiforme. *J Clin Oncol* 2000, **18**, 1481–1491.
37. Ware Jr JE, Bjorner JB, Kosinski M. Practical implications of item response theory and computerized adaptive testing: a brief summary of ongoing studies of widely used headache impact scales. *Med Care* 2000, **38**(Suppl. 9), 73–82.
38. Cella D, Chang CH. A discussion of item response theory and its applications in health status assessment. *Med Care* 2000, **38**(Suppl. 9), 66–72.