PERGAMON

European Journal of Cancer 37 (2001) 1081–1088

European Journal of Cancer

www.ejconline.com

# Deriving a compound quality of life measure from the EORTC-QLQ-C30/LC13 instrument for use in economic evaluations of lung cancer clinical trials

A. Bagust *, M. Barraza-Llorens, Z. Philips

*York Health Economics Consortium, University of York, Heslington, York YO10 5EL, UK*

Received 20 September 2000; received in revised form 5 January 2001; accepted 9 February 2001

**Abstract**

Many clinical trials involve parallel collection of quality of life (QoL) and economic data, requiring patients to complete similar questionnaires at regular intervals. This additional burden often leads to disappointing response rates and inconclusive results. Data obtained in the LU-16 trial with the European Organization for Research and Treatment of Cancer Quality of Life (EORTC-QLQ-C30)/LC-13 QoL instrument for lung cancer were re-analysed, using multivariate techniques. The analysis demonstrated the inherent non-linearity of QoL data, with resulting interpretational problems. A new integrated linear QoL measure was developed which maximises the use of the information collected and can serve as a proxy utility measure for economic evaluation. It was successfully validated with data from another lung cancer trial with encouraging results. For individual patients, trends in QoL are revealed more clearly with narrower confidence intervals. This approach yields relative weightings and rankings for the main issues affecting QoL ratings in lung cancer patients, most importantly fatigue, breathlessness, poor concentration and disruption to family and social life. 

*Keywords:* Quality of life; Economics; Multivariate analysis

## 1. Introduction

Quality of Life (QoL) as a health-related outcome is widely recognised as an important component of both clinical trials and economic evaluations. The assessment of QoL changes is of particular importance in lung cancer (whether or not predictive of survival [1,2]) where over 80% of patients die within a year of diagnosis, and treatment tends to be largely palliative [3,4].

There is no single agreed definition of QoL and various approaches have been adopted for its measurement. However, several multidimensional self-rated instruments have been developed in accord with the prevailing consensus [3,5–7] resulting in the demotion of earlier uni-dimensional functional measures (e.g. Karnofsky Performance Scale) [8]. The development of these QoL instruments has focused on the inclusion of relevant dimensions or domains that are deemed to determine QoL, principally physical functioning (including mobility), psychological functioning (including emotional and cognitive well-being) and social functioning [9].

From the clinical perspective, QoL instruments should measure items with known relevance for the patient–disease groups under study. This further detail is provided by the development of specific instruments more responsive to treatment-related changes for a particular disease [10]. However, the need for comparability in health-related QoL outcomes between patient groups still requires instruments of sufficient generality to allow their use in different clinical contexts. As a result, instruments have been developed which measure general dimensions of QoL, but also cover elements relevant for specific patients, diseases or treatments through complementary questionnaires (e.g. European Organization for Research and Treatment of Cancer Quality of Life Core 30 questionnaire (EORTC-QLQ-C30) and its lung cancer module (LC-13) [11], and functional assessment of cancer therapy-general (FACT-G) and its lung cancer complement [10,12–14]). As instruments have gained a greater degree of refinement and specificity, the published

* Corresponding author. Tel.: +44-1904-433620; fax: +44-1904-433628.

*E-mail address:* ab13@york.ac.uk (A. Bagust).

0959-8049/01/$ - see front matter 
PII: S0959-8049(01)00078-8

literature has concentrated mainly on their psychometric properties (validity and reliability) and on ensuring that all relevant domains are taken into account, including symptoms and side-effects specific not only to the disease, but to therapeutic interventions. However, the quest for comprehensive and reliable measures through multidimensionality does not necessarily aid interpretation.

By contrast, economic evaluators and health policy analysts require QoL assessments which measure items relevant to specific patient groups but which may be compared between different patient groups. In addition, aggregation of all the QoL information collected into a single, meaningful measure of health-related QoL is essential to facilitate cost-effectiveness or cost-utility analysis where QoL is the primary health outcome [9,14–16]. Some of the instruments currently available contain single scale ratings of patients' assessment of their overall QoL. Alternatively, total QoL scores may be constructed by adding or averaging scores from different scales. Generally, these attempts to provide a single-measure are simplistic and researchers often fail to explore the interrelationship between domains and overall scores except at a rudimentary level. Consequently, there remains a gap between the uses to which the bulk of information collected through QoL instruments is put and the availability and deployment of a single overall measure of QoL suitable for economic evaluations [14–17]. Although QoL is undoubtedly complex, both conceptually and practically, decision-makers have a legitimate expectation that QoL instruments yield clear and interpretable results.

At present, QoL and health economics are often treated as disjoint elements of trial design, when in reality they share common needs for evidence on outcomes only available from patients and carers. However, when QoL and health economic data needs are not co-ordinated, an excessive burden may be imposed on vulnerable patients. The consequence is likely to be reduced compliance with all data collection instruments leading to inadequate or ambiguous results. If re-interpretation of data from existing instruments can be made to serve both purposes acceptably, there will be benefits for patients and for researchers of all persuasions.

The EORTC-QLQ-C30-LC13 [11] has proven to be a valid and reliable psychometric instrument [18,19] to assess health-related QoL as an outcome measure in clinical trials of patients with lung cancer. Besides a range of detailed items, it includes two Likert scales (questions 29 and 30) to capture a patient's assessment of physical condition and overall quality of life. This study was designed to investigate the feasibility of exploiting the full range of detailed information the instrument elicits to obtain a meaningful overall QoL measure useful for economic evaluations and which may be more robust and sensitive than the original Likert scales. This involves applying quantitative analytical procedures in a novel way to explore the underlying relationships between QoL domains and overall QoL scores recorded in the questionnaire. If successful, this would allow researchers to proxy the measurement of patient-based utility within the context of cost-utility analysis without additional data collection from patients.

## 2. Patients and methods

### 2.1. Main data set

The principal data set for model development comprised 330 EORTC QLQ-C30-LC13 responses completed by 91 small-cell lung cancer (SCLC) patients as part of the MRC LU-16 trial of palliative chemotherapy [20]. The demographic profile of patients and responses is shown in Table 1. Forms were completed at randomisation and once per 3-weekly cycle. The number of forms completed by patients varied between one (26 patients) and 10 (2 patients) with a mean of 3.6 per patient, reflecting progressive loss of participation through mortality and non-response to the questionnaire. The clinical characteristics of patients at randomisation were similar to those for the whole trial: 34 patients (38%) were classified with limited disease, and 55 (62%) with extensive disease (2 unknown); initial performance status was rated as moderate in 56 patients (62%), poor in 33 (36%) or very poor in 2 (2%).

In the completed forms, questions in the common core (numbers 1–31) exhibited very high response rates, averaging 99.2% with no question achieving less than 97.9%. The supplementary lung cancer section (LC13), was slightly less successful with an average completion rate of 96.1% and the worst question answered on only 90% of forms (question 42 'pain in other parts of body'').

Table 1
Demographic profile of LU-16 patients completing EORTC forms

| Gender | Age | Patients *n* (%) | Forms *n* (%) |
|---|---|---|---|
| Male | Under 40 | 1 (1) | 5 (1.5) |
| | 40–49 | 2 (2) | 2 (0.6) |
| | 50–59 | 8 (9) | 39 (11.8) |
| | 60–69 | 29 (32) | 89 (27.0) |
| | 70 and over | 19 (21) | 83 (25.2) |
| Female | Under 40 | 0 (0) | 0 (0.0) |
| | 40–49 | 0 (0) | 0 (0.0) |
| | 50–59 | 6 (7) | 31 (9.4) |
| | 60–69 | 12 (13) | 49 (14.8) |
| | 70 and over | 13 (14) | 31 (9.4) |
| Unknown | Missing | 1 (1) | 1 (0.3) |
| | Total | 91 (100) | 330 (100.0) |

EORTC, European Organization for Research and Treatment of Cancer.

Table 2
Demographic profile of JHDN patients completing the EORTC forms

| Gender | Age (years) | Patients *n* (%) | Forms (%) |
|---|---|---|---|
| Male | Under 40 | 1 (0.3) | 6 (0.4) |
| | 40–49 | 9 (3.1) | 46 (3.3) |
| | 50–59 | 41 (13.9) | 159 (11.3) |
| | 60–69 | 87 (29.5) | 457 (32.5) |
| | 70 and over | 51 (17.3) | 229 (16.3) |
| Female | Under 40 | 2 (0.7) | 9 (0.6) |
| | 40–049 | 5 (1.7) | 19 (1.4) |
| | 50–59 | 20 (6.8) | 104 (7.4) |
| | 60–69 | 49 (16.6) | 218 (15.5) |
| | 70 and over | 30 (10.2) | 157 (11.2) |
| | Total | 295 (100.0) | 1404 (100.0) |

EORTC, European Organization for Research and Treatment of Cancer.

### 2.2. Data analysis and statistical modelling

Two preparatory stages of data manipulation preceded the main analysis and statistical modelling. First, it was determined that both detailed items and the overall scales in the EORTC instrument exhibited pronounced non-linearity in responses, sufficient to compromise subsequent multivariate analyses. A technique of maximising interitem correlations was used to rescale the responses to each question to reduce this effect (see Appendix, Section A1).

Second, the presence of several questions with overlapping domains threatened to destabilise the performance of the multiple linear regression in the statistical modelling. Therefore, a set of new combined variables was developed in order to reduce intervariable collinearity to a minimum among the detailed questionnaire items (Appendix, Section A2).

The main statistical modelling involved an iterative process using both principal components analysis and multiple linear regression to represent the overall ratings of patients (questions 29 and 30) in terms of the detailed items (Appendix, Section A3).

### 2.3. Validation

The validity of the models was tested by applying the models to EORTC data obtained as part of a trial of Gemcitabine in the treatment of advanced non-small cell lung cancer (NSCLC) (JHDN). This provided 1404 EORTC records from 295 patients at approximately 4-weekly intervals over 6 months, with an age/sex distribution (Table 2) very similar to that of the LU-16 patients. Correspondence between model estimates and patient ratings was assessed for both sets of data by calculation of intraclass correlations, conventional Kappa coefficients, and a modified Kappa coefficient using the linearising interval weights for the overall scales. Table 3 shows generally acceptable performance for both models, and no loss of precision between the two data sets.

## 3. Results

In the process of developing the Physical Condition model (question 29), it proved necessary to discard four variables — sore mouth (S), tingling (U), hair loss (V) and financial difficulties (P). These were similarly excluded from the Quality of Life model (question 30), together with pain control (W), constipation (I) and memory (N). The resulting models accounted for 57 and 59% respectively of variance in the standardised scales. Model residuals were subjected to analysis of variance, and found to be unrelated to disease stage, performance status, age or familiarity with the EORTC instrument. However, there was a significant difference ($P<0.05$) for gender, with men rating higher than women by 5.9 on the Physical Condition scale, and 5.6 on the Quality of Life scale.

By combining principal metric weightings with linear regression coefficients, estimates were obtained of the contributions of each standardised variable to the overall estimates of Physical Condition and Quality of Life; these are shown graphically in Fig. 1. Very similar patterns are obtained for the two scales, suggesting that there is a large degree of communality in response to questions 29 and 30. The most important factor is clearly fatigue, followed by breathlessness, poor concentration and disruption to family and social life. The significance of the differences between weightings in the different models cannot be readily quantified, but it appears that seven variables are rated as markedly more influential on patients' perception of overall QoL, than

Table 3
Correspondence test statistics for the original and validation data sets

| Correspondence statistic | Question 29 — Physical condition | | Question 30 — Quality of life | |
|---|---|---|---|---|
| | LU-16 (SCLC) | JHDN (NSCLC) | LU-16 (SCLC) | JHDN (NSCLC) |
| Intraclass correlation | 0.67 | 0.67 | 0.68 | 0.67 |
| Simple Kappa | 0.46 | 0.48 | 0.49 | 0.47 |
| Modified Kappa | 0.51 | 0.51 | 0.53 | 0.50 |
| No. of records | 329 | 1182 | 327 | 1221 |

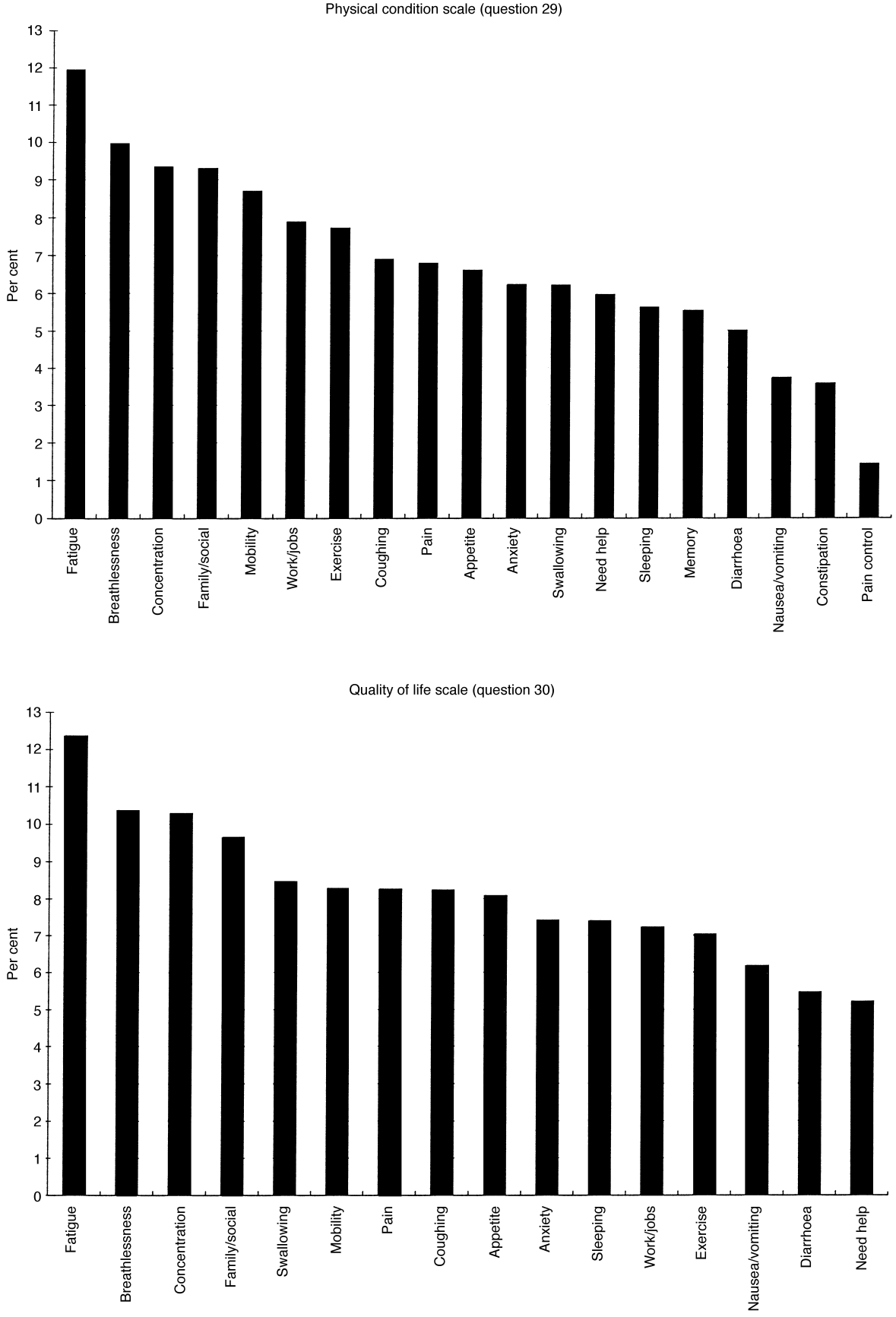


Fig. 1. Model component weightings.

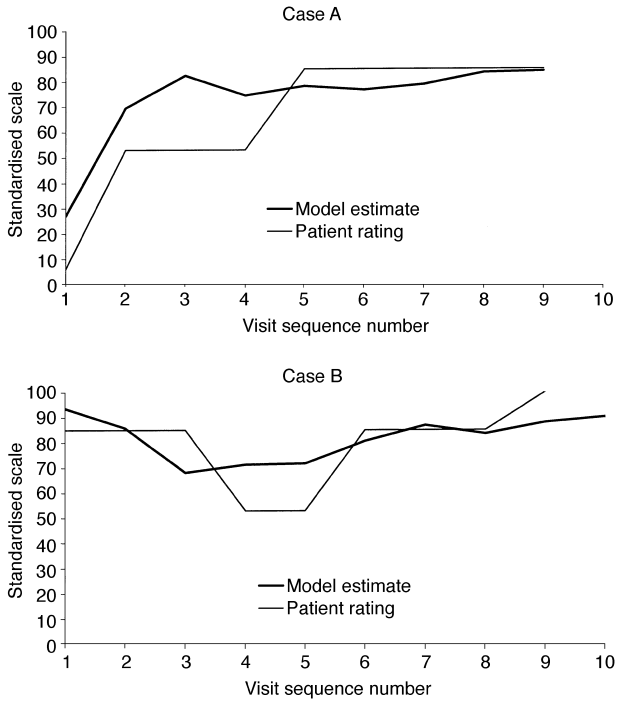


Fig. 2. Comparison of QoL model scores and EORTC question 30 for 2 patients.

on physical condition: nausea–vomiting (H), swallowing (T), sleeping (F), pain (K), appetite (G), coughing (Q) and anxiety (M).

Fig. 2 illustrates the use of the model estimates in tracking QoL changes over time, compared with the corresponding Likert-scale rating. For 2 JHDN patients completing nine or 10 EORTC-QLQ-C30/LC-13 questionnaires, the models avoid large step changes and generally give a clearer indication of the QoL trend.

The size of differences in quantified QoL scores which show statistical significance depends critically on the population variance of the measured variable. The two models generate estimates which are more nearly continuous functions than questions 29 and 30 (restricted to integer steps). It would therefore be expected for these to exhibit reduced variance and hence improved precision. This proved to be the case, with variances reduced by 23 and 24.5%, respectively, equivalent to 12–13% narrower confidence intervals (CI). When considering the distribution of differences between adjacent paired observations, the variance reductions were even greater (41 and 36% — giving 23 and 20% narrower CI).

## 4. Discussion

The approach adopted in this study has demonstrated for the first time the feasibility of deriving an effective and reproducible utility measure from information collected from use of a QoL questionnaire. This offers the prospect of integrating QoL and health economics components of clinical trials, with reduced data collection costs and fewer burdens on patients. In addition, it allows researchers and decision-makers to understand the relative importance of particular symptoms to patients, offering some remedy to the lack of interpretability of both existing QoL instruments and generic economic utility measures. By maximising the use of information contained in the whole EORTC data set, the precision of inference testing in a clinical trial is also enhanced, improving the prospects of obtaining meaningful QoL and economic findings from samples designed for conventional clinical outcomes. The training data set (LU-16) was restricted in size and should be supplemented with larger studies where possible to confirm the robustness of the analytical approach. However, the extent of confirmation provided by the larger validation set (JHDN) gives some confidence that the results reported probably reflect genuine effects.

### 4.1. Quality of life instruments

The re-analysis of QoL data was not intended to address any issues in relation to the methodology of psychometric scale development, nor to comment on the integrity of the EORTC instrument specifically. However, some conclusions can be drawn on the basis of this study.

We have demonstrated the inherent non-linearity of most elements and sub-scales of the QoL instrument and presumably of others currently available. This suggests that they are not suitable for utility analysis without modification since this requires interval equivalence throughout a scale. It also suggests that tests for statistical significance used in clinical trials should be modified for differences in variance across the range of any QoL sub-scale.

Item redundancy is inherent in psychometric scale construction in order to achieve scale reliability. We have demonstrated the scale of such redundancy (seven out of 41 candidate items were not present in either model) but, more significantly, revealed the extent to which alternative combination questions might more accurately capture the nuances of patient response (25 items combined into 10 new variables).

Relative weightings derived for each questionnaire item are novel in revealing the variation in the magnitude of these effects, although the most significant factors have previously been reported (e.g. fatigue [21]). It is notable that, although separate models were developed for physical condition (question 29) and overall quality of life (question 30), the ranking and weighting of items show few differences, suggesting that patients may have difficulty in distinguishing between the two scales in practice. Where differences are apparent, it is notable that the impact of physical symptoms is not restricted to the physical condition scale and, in fact, in

several instances overall QoL is more seriously affected by severe symptoms (nausea/vomiting, swallowing, sleeping, pain, appetite and coughing). It may be argued that with time and familiarity the perception of some symptoms *per se* may lessen, particularly where interventions or coping strategies are brought into play. Yet the functional limitations such symptoms impose on social interaction remain an impediment to normal life.

### 4.2. Economic evaluation

It is clear that the modelled QoL scales developed do not address all the questions raised by health economists. Most notably, they do not resolve the problem of explicit state valuation which economic theory requires in addition to health status definition in order to determine patients' willingness to trade-off between the utility/disutility attached to any particular health state. Whether patients would prefer feeling more fatigue but less anxiety is an issue that ultimately will determine how patients value particular health states and therefore treatment outcomes.

In this study, patients' preferences are not explicitly elicited as in the case of other instruments (e.g. EuroQoL), but the explanatory role of each item in the overall QoL observed for a patient may indicate indirectly which elements are key in determining patients' perception of their overall QoL. Consequently, the models can be considered an implicit approximation for patients' preferences, at least as far as the EORTC instrument covers the issues relevant to patients' valuations.

Although promising as a potential utility measure, the model scores we have developed here only have any applicability for lung cancer patients and require explicit calibration against a recognised generic utility measure such as EuroQoL EQ-5D. This would facilitate comparisons of outcomes and utilities with other patient–disease groups for health policy analysis, and also allow conversion of the model scores into cardinal measures by establishing their relationship to absolute states ('death' and 'full health'). With this enhancement, the re-interpretation of data from QoL questionnaires could be directly integrated into economic evaluations of clinical trials, significantly reducing the data collection burden on both patients and researchers.

### 4.3. Instrument design and understanding

The divergence of approach between QoL researchers and health economists illustrates their radically different evaluative paradigms. For QoL practitioners, a key criterion for success is the ability to reflect the multidimensional nature of human perception and experience. By contrast, the economists' imperative to achieve a common component of personal utility has led to techniques designed to subsume diversity within a single unified measure. Clinical trials are a key means by which clinical practice and health policy generally are informed but, sadly, too little attention has been paid to interpretability in the design and reporting of results. The application of a novel approach to the analysis of QoL data has enabled us to relate a general utility measure suitable for economic analysis to details of a QoL instrument and this offers the prospect of providing more meaningful evidence to decision-makers from clinical trials to the benefit of clinical researchers, QoL specialists and health economists alike.

## Acknowledgements

The authors are grateful to Dr P. Hopwood and the Molecular Research Council (MRC) for permission to use the LU-16 data. This research was supported by grant from Eli Lilly and Company Ltd.

## Appendix. Analytical methodology

### A1. Linearising question scales

In questions 29 and 30 of EORTC-QLQ-C30, patients are asked to rate their overall physical condition (question 29) and quality of life (question 30) during the preceding week on a seven-point scale from 'Very poor' (scored 1) to 'Excellent' (scored 7). The objective of the analysis was to use combinations of responses made by patients to the detailed questions on the form in order to model their responses to questions 29 and 30. However, since all the questionnaire items yielded ordered categories rather than continuous scale measures, it is not clear whether any of the information obtained can be treated as essentially linear, as is necessary in order to carry out economic evaluations. If any item is inherently non-linear (i.e. a change of one interval step at the bottom of the scale cannot be equated to one step elsewhere on the scale) then it is important to correct for such behaviour before further analysis is attempted.

Individual questions (excluding those with binary responses — numbers 1–7), and the EORTC sub-scale scores were plotted against patients' overall status estimates (questions 29 and 30). The results revealed a varied picture in which some questions and sub-scales appeared to change in a broadly linear pattern compared with the overall measures, whereas others exhibited markedly non-linear behaviour, often of a sigmoid (S) shape. Because there was no single general pattern affecting all the detailed questions and sub-scales similarly, there was no evidence that non-linear behaviour was solely due to the overall status scales (questions 29 and 30), and therefore it was concluded that features of

individual questions were an important source of the observed non-linearity.

The basis of linearising question responses was to adjust the relative size of the scale intervals (1–2, 2–3, 3–4) for a question in order to maximise the Spearman correlation with the overall status questions (29 and 30), thus obtaining the closest approximation possible to a linear relationship. The responses to each question were then rebased on a scale from 0 (best) to 100 (worst), as shown in Table A1; the extent to which the revised scores do not increase in equal steps reveals the degree of inherent non-linearity in the individual question ratings.

Following these transformations made to the detailed EORTC items, it was necessary to consider whether a similar non-linearity remained unresolved in the overall status questions (29 and 30). To address this question, we created a new variable as an average combination of the linearised detailed items, and then used a similar procedure to rescale the Likert scale intervals on questions 29 and 30 so as to maximise their correlations with the new variable. The resulting values are shown in Table A2 and it is clear that the extreme values are poorly distinguished by most patients. In view of the similarity of the weighting schemes derived for questions 29 and 30 and the limitations of a relatively small sample, it was decided to adopt a single average scheme for use with both overall items in subsequent analyses.

### *A2. Combination and redundancy of questions*

In the development of psychometric instruments, closely related questions are often included to provide stability to sub-scales. However, this may result in considerable overlap between items, and therefore possible redundancy within data gathered with the instrument, evidenced by high inter-item correlations. This creates problems when using some multivariate statistical analysis techniques (e.g. regression), resulting in unstable and unreliable results. To avoid such difficulties, pairs or groups of questions with similar content and Spearman correlation coefficients greater than 0.4 were identified. These were replaced by new combination variables, which encompassed the information within each original set of questions, but reduced collinearity within the whole data set. The new combination variables are shown in Table A3, including 12 which

Table A1
Rescaled values of individual EORTC question responses

| Question no. | Original responses | | | | Question no. | Original responses | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | | 0 | 1 | 2 | 3 | 4 |
| 8 | 0 | 28 | 65 | 100 | 25 | – | 0 | 39 | 83 | 100 |
| 9 | 0 | 29 | 75 | 100 | 26 | – | 0 | 22 | 76 | 100 |
| 10 | 0 | 13 | 64 | 100 | 27 | – | 0 | 0 | 58 | 100 |
| 11 | 0 | 12 | 56 | 100 | 28 | – | 0 | 13 | 100 | 100 |
| 12 | 0 | 24 | 68 | 100 | 31 | – | 0 | 0 | 42 | 100 |
| 13 | 0 | 50 | 70 | 100 | 32 | – | 0 | 50 | 100 | 100 |
| 14 | 0 | 36 | 36 | 100 | 33 | – | 0 | 47 | 65 | 100 |
| 15 | 0 | 60 | 100 | 100 | 34 | – | 0 | 24 | 67 | 100 |
| 16 | 0 | 34 | 34 | 100 | 35 | – | 0 | 34 | 57 | 100 |
| 17 | 0 | 47 | 47 | 100 | 36 | – | 0 | 56 | 100 | 100 |
| 18 | 0 | 14 | 70 | 100 | 37 | – | 0 | 46 | 100 | 100 |
| 19 | 0 | 30 | 94 | 100 | 38 | – | 0 | 5 | 5 | 100 |
| 20 | 0 | 35 | 62 | 100 | 39 | – | 0 | 81 | 88 | 100 |
| 21 | 0 | 33 | 78 | 100 | 40 | – | 0 | 26 | 80 | 100 |
| 22 | 0 | 19 | 61 | 100 | 41 | – | 0 | 7 | 40 | 100 |
| 23 | 0 | 22 | 100 | 100 | 42 | – | 0 | 30 | 41 | 100 |
| 24 | 0 | 39 | 93 | 100 | 43/44 | 0 | 10 | 74 | 100 | 100 |

EORTC, European Organization for Research and Treatment of Cancer.

Table A2
Rescaled linearised values assigned to overall status questions

| | Original response | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Question 29 | 0 | 11 | 38 | 55 | 84 | 100 | 100 |
| Question 30 | 0 | 0 | 29 | 50 | 87 | 100 | 100 |
| Average | 0 | 6 | 34 | 53 | 85 | 100 | 100 |

Table A3
Standardised linearised variables selected for modelling

| New variable | Original questions | Issue addressed | Combination method |
|---|---|---|---|
| A | 1,2 | Strenuous exercise | Simple average |
| B | 3,4 | Mobility | Simple average |
| C | 5 | Needing help | None |
| D | 6,7 | Work/jobs limitations | Combination with question 7 dominant |
| E | 10,12,18 | Fatigue | Simple average |
| F | 11 | Sleeping | None |
| G | 13 | Appetite | None |
| H | 14,15 | Nausea and vomiting | Maximum reweighted score |
| I | 16 | Constipation | None |
| J | 17 | Diarrhoea | None |
| K | 19,40,41,42 | Experience of pain | Maximum reweighted score |
| L | 20 | Concentration | None |
| M | 21,22,23,24 | Anxiety | Maximum reweighted score |
| N | 25 | Memory | None |
| O | 26,27 | Family and social life | Simple average |
| P | 28 | Financial problems | None |
| Q | 31,32 | Coughing | Maximum reweighted score |
| R | 34,35 | Breathlessness | Simple average |
| S | 36 | Sore mouth | None |
| T | 37 | Swallowing | None |
| U | 38 | Tingling | None |
| V | 39 | Hair loss | None |
| W | 43/44 | Pain control | None (43 and 44 form a single variable) |

remained uncombined, and three which were found to be effectively redundant (adding no additional information to that associated with other items). In total, 23 new standardised variables were identified as suitable for the final stages of the analysis.

### *A3. Multivariate modelling*

The modelling objective was to develop representations of the general assessment questions (29 and 30 reweighted) in terms of linear combinations of the new standardised and linearised variables derived from the detailed EORTC questionnaire items. These should account for as much as possible of the variance in the data sample but should avoid non-intuitive relationships generated as artefacts of the analytical techniques (usually as a result of collinearity between variables).

A two-stage iterative procedure was carried out. Firstly, orthogonal principal components analysis was used with the 23 standardised variables to generate new weighted combination variables (principal components) which account for the largest possible proportion of the variance within the data. Only principal components which perform better than a chance combination (i.e. with eigenvalues greater than 1) were selected. Then, these were used as independent variables in a multiple linear regression analysis to model questions 29 and 30.

Where this procedure yielded model coefficient(s) with counter-intuitive signs, the offending variables were excluded one at a time, and the whole two-stage process was repeated. This iterative process was continued until all the remaining standardised variables conformed with expected patterns.

## References


1. Herndon II JE, Fleishman S, Kornblith AB, Kosty M, Green MR, Holland J. Is quality of life predictive of the survival of patients with advanced nonsmall cell lung carcinoma? *Cancer* 1999, **85**, 333–340.
2. Dancey J, Zee B, Osoba D, *et al*. Quality of life scores: an independent prognostic variable in a general population of cancer patients receiving chemotherapy. *Qual Life Res* 1997, **6**, 151–158.
3. Montazeri A, Gillis CR, McEwen J. Quality of life in patients with lung cancer: a review of literature from 1970 to 1995. *Chest* 1998, **113**, 467–481.
4. Kosmidis P. Quality of life as a new end point. *Chest* 1996, **109**(Suppl.), 110S–112S.
5. Berzon RA, Donnelly MA, Simpson RL, Simeon GP, Tilson HH. Quality of life bibliography and indexes: 1994 update. *Qual Life Res* 1995, **4**, 547–569.
6. Kassa S, Mastekaasa A, Naess S. Quality of life of lung cancer patients in a randomized clinical trial evaluated by a psychosocial well-being questionnaire. *Acta Oncol* 1998, **27**, 335–342.
7. Montazeri A, Gillis CR, McEwen J. Measuring quality of life in oncology: is it worthwhile I. Meaning, purposes and controversies? *Eur J Cancer Care* 1996, **5**, 159–167.
8. Schaafsma J, Osoba D. The Karnofsky Performance status scale re-examined: a cross-validation with the EORTC-C30. *Qual Life Res* 1994, **3**, 413–424.
9. Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-effectiveness in health and medicine*. New York, Oxford University Press, 1996.
10. Sprangers MAG, Cull A, Bjordal K, Groenvold M, Aaronson NK. The European Organization for Research and Treatment of Cancer approach to quality of life assessment: guidelines for developing questionnaire modules. *Qual Life Res* 1993, **2**, 287–295.
11. Bergman B, Aaronson NK, Ahmedzai S, Kaasa S, Sullivan M. The EORTC QLQ-LC13: a modular supplement to the EORTC core quality of life questionnaire (QLQ-C30) for use in lung cancer clinical trials. EORTC study group on quality of life. *Eur J Cancer* 1994, **30A**, 635–642.
12. Sprangers MAG, Cull A, Groenvold M, Bjordal K, Blazeby J, Aaronson NK. The European Organization for Research and Treatment of Cancer approach to developing questionnaire modules: an update and overview. *Qual Life Res* 1998, **7**, 291–300.
13. Cella D. The functional assessment of cancer therapy-anemia (FACT-An) scale: a new tool for the assessment of outcomes in cancer anaemia and fatigue. *Semin Hematol* 1997, **34**(Suppl. 2), 13–19.
14. Cramer JA, Spilker B. *Quality of Life and Pharmacoeconomics. An Introduction*. Philadelphia, Lippincott-Raven Publishers, 1998.
15. Torrance GW. Designing and conducting cost-utility analyses. In Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. Philadelphia, Lippincott-Raven Publishers, 1996, 1105–1111.
16. Revicki DA. Relationship between pharmaeconomics and health related quality of life. In Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. Philadelphia, Lippincott-Raven Publishers, 1996, 1077–1083.
17. Aaronson NK, Ahmedzai S, Bergman B, *et al*. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993, **85**, 365–376.
18. Osoba D, Zee B, Pater J, Warr D, Kaizer L, Latreille J. 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**, 353–364.
19. Hollen PJ, Gralla RJ. Comparison of instruments for measuring quality of life in patients with lung cancer. *Semin Oncol* 1996, **23**(Suppl. 5), 31–40.
20. MRC Lung Cancer Working Party. Comparison of oral etoposide and standard intravenous multidrug chemotherapy for small-cell lung cancer: a stopped multicentre randomised trial. *Lancet* 1996, **348**, 563–566.
21. Sarna L, Brecht M. Dimensions of symptom distress in women with advances lung cancer: a factor analysis. *Heart Lung* 1997, **26**, 23–30.