Derivation of Utility Values from European Organization for Research and Treatment of Cancer Quality of Life-Core 30 Questionnaire Values in Lung Cancer

Raymond W. Jang, MD,† Pierre K. Isogai, BSc,‡ Nicole Mittmann, PhD,‡§ Penelope A. Bradbury, MD,* Frances A. Shepherd, MD,*† Ronald Feld, MD,*† and Natasha B. Leighl, MD*†*

Introduction: Cancer clinical trials frequently incorporate quality of life (QoL) measures but rarely patient utility. Utility information is required for cost utility evaluations of novel cancer therapies. We assessed the feasibility of converting QoL data into utility scores using the European Organization for Research and Treatment of Cancer Quality of Life-Core 30 questionnaire (EORTC QLQ-C30) and the EQ-5D in patients with non-small cell lung cancer (NSCLC).

Methods: Outpatients with all different disease states of NSCLC attending a major Canadian cancer center completed the QLQ-C30 and EQ5D on a single visit. Results of the QLQ-C30 summary scores were mapped to predict EQ-5D utility scores using linear regression. Backward variable elimination using the Akaike Information Criterion was used to reduce the full model that included all QLQ-C30 summary scores to examine which QLQ-C30 dimensions best predict a patient's utility score. To test the predictive power of the model, 10-fold cross-validation was used.

Results: A total of 172 patients participated in the study. Median age of the sample was 66 years (range, 32–85 years); 46.5% were men. The cross-validation estimate of mean utility score was 0.76 (SD: 0.20), which was the same as the actual mean utility score. Of the 15 QLQ-C30 dimensions, 4 functional dimensions (physical, role, emotional, and social) and the pain symptom dimension were predictive of patient utility scores.

Conclusions: Our study demonstrates the feasibility of deriving utility scores from prospective QoL data. Validation of the QLQ-C30 predictors found in this study could further the ability to estimate cost utility of therapies for economic evaluations.

Key Words: Non-small cell lung cancer, EQ-5D, EORTC QLQ-C30, Derivation of utility values, Economic evaluation.

(*J Thorac Oncol.* 2010;5: 1953–1957)

Cancer clinical trials frequently measure quality of life (QoL) using multidimensional QoL instruments, such as the European Organization for Research and Treatment of Cancer Quality of Life Core 30 questionnaire (EORTC QLQ-C30)[1] and the Functional Assessment of Cancer Therapy-General (FACT-G) questionnaire.[2] The questionnaires yield multiple values corresponding to different dimensions of patients' QoL; however, in economic analyses such as cost-utility analyses, a single universal measure of health outcome, such as a utility measure, is needed to compare different therapies.[3] In other words, utility measures differ from QoL measures in that they combine the positive and negative aspects of a health state into a single number and thereby provide a common unit of analysis when comparing different types of treatment.[3] Patient utility is commonly used as a measure of preference, anchored at 0 (death) and 1 (perfect health).[4] Despite the preference for utility measurements, most clinical trials do not collect data routinely on patient utility, in part because of time constraints, costs, and lack of interest from sponsors and investigators.

When utility values are required for economic evaluations of novel therapies but have not been collected as part of clinical trials, there is a need to consider alternative approaches. Alternate approaches to the prospective collection of utilities include the use of a global QoL question incorporated in some QoL instruments or the use of linear scoring rules wherein subscales of different dimensions are summed to produce a comprehensive score. However, none of these approaches produces a true utility value that corresponds to an established decision science framework, because true utility scores should be based on preferences for different health states.[4] Furthermore, linear scoring rules may be mis-

*Division of Medical Oncology, Princess Margaret Hospital/University Health Network; †Department of Medicine, University of Toronto; ‡Health Outcomes and Pharmacoeconomic Evaluation Research Centre, Sunnybrook Health Sciences Centre; and §Department of Pharmacology, University of Toronto, Toronto, Ontario, Canada.

Disclosure: The authors declare no conflict of interest.

Address for correspondence: Natasha Leighl, MD, MMSc, FRCPC, Princess Margaret Hospital, Suite 5-105, 610 University Avenue, Toronto, ON, Canada M5G 2M9. E-mail natasha.leighl@uhn.on.ca

Presented in part at the 13th Annual Meeting of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), Toronto, 2008, and the 13th World Conference on Lung Cancer, San Francisco, 2009.

Jang and Isogai contributed equally to this article. Jang received a Young Investigator Travel Award at the 13th World Conference on Lung Cancer.

ISSN: 1556-0864/10/0512-1953

leading because of nonlinearity in preferences for different outcomes.[5] For these reasons, cost utility evaluations of novel cancer therapies remain a challenge. A third alternative, where utility data have not been collected prospectively in a trial, is to use statistical models to convert prospectively collected QoL data into utility scores. Mortimer has reviewed these models, some of which have been developed for patients with cancer.[6,7] Our study used this third alternative approach and explored the association between QoL data obtained with two commonly used instruments, the EORTC QLQ-C30 and the EQ-5D, from patients with non-small cell lung cancer (NSCLC).

METHODS

A total of 172 consecutive outpatients with NSCLC attending a major Canadian cancer center outpatient clinic were invited to complete the EORTC QLQ-C30 and EQ-5D at a single visit. Demographic information including age, disease stage, and disease state were extracted from patient charts. Disease states included (1) relapse-free post-resection, including on adjuvant chemotherapy; (2) in relapse but on no treatment or receiving palliative chemotherapy or targeted therapy (erlotinib); and (3) receiving radical chemoradiation for locally advanced (stage III) NSCLC.

The EORTC QLQ-C30 is a self-administered, cancer-specific questionnaire that has multidimensional scales consisting of five functional domains (physical, role, emotional, cognitive, and social); three symptom domains (fatigue, nausea/vomiting, and pain); six single-item symptom scores (dyspnea, sleep, appetite, constipation, diarrhea, and financial impact), and a global health/QoL domain (overall health and overall QoL in the past week).[1] Responses are scored on a scale from 1 (not at all) to 4 (very much), with the exception of questions related to the global QoL domain scale, which are scored on a scale from 1 (very poor) to 7 (excellent). For our analysis, the 30 questions of the QLQ-C30 were collapsed into 15 dimensions listed above as per the EORTC QLQ-C30 scoring manual.[8] QLQ-C30 scores for these variables were obtained by first calculating the raw score (RS), where $RS = (I_1 + \ldots + I_n)/n$, where I_1 to I_n are the responses to questions included in the variable. Then a linear transformation was applied to scale the scores between 0 and 100. For functional and symptom scores, the new score $= (1 - [\{RS - 1\}/n]) \times 100$, and for the global QoL score, the new score $= ([RS - 1]/n) \times 100$. Thus, a high score for a functional or global dimension represents a high level of functioning, but a high score for a symptom dimension represents a high level of symptomatology. Finally, the QLQ-C30 does not provide a direct or indirect measure of utility.

The EQ-5D is a generic health-related QoL instrument that consists of five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.[9,10] Because each dimension has three possible levels, 243 (3^5) unique health states exist. It also includes a visual analog scale for self-assessment of current general health. The EQ-5D provides an indirect measure of utility scores, because valuation studies with different populations have been done where each possible health state in EQ-5D has been assigned a utility score using validated methods such as the time trade-off method. The end result of these valuation studies is a schedule of utility scores or tariffs corresponding to different health states for a specific population, usually a country. For our study, the US population tariffs were used.[11]

Results of the QLQ-C30 were mapped to predict EQ-5D utility scores using a linear regression model. Utility scores obtained from the EQ-5D were regressed on the 15 QLQ-C30 dimensions. Backward variable elimination[12] using the Akaike Information Criterion (AIC) was used to reduce the full model to examine which QLQ-C30 dimensions best predict a patient's utility score. AIC was used to account for the tradeoff between complexity and fit of the model. Coefficients for the linear regression model were calculated for the full model and the reduced model. Interactions between variables were not investigated.

To test the predictive performance, 10-fold cross-validation was used. In 10-fold cross-validation, the sample is partitioned into 10 approximately equal subsets. Models were fit using nine of the subsets, and the remaining subset is used to validate the fitted model. The process was completed 10 times, such that each subset was used as the validation sample exactly once. Mean squared error was calculated to measure deviation of the predicted from the actual utility value. Predicted and actual utility scores were also compared by disease state and stage. A 3000-replicate bootstrap was used to derive 95% confidence intervals (CIs). All analyses were performed using the R language for statistical computing, version 2.8.

The protocol was approved by the research ethics board at the Princess Margaret Hospital, Toronto, Canada, and all patients provided written informed consent to participate.

RESULTS

The characteristics of participants are listed in Table 1. All four stages of NSCLC were represented, although the

TABLE 1. Characteristics of Participants ($n = 172$)

Age (yr)	
Median	66
Range	32–85
Male sex (%)	80 (46.5)
Disease stage (%)	
Stage I	34 (19.8)
Stage II	16 (9.3)
Stage III	36 (20.9)
Stage IV	86 (50)
Clinical state	
Relapse free	
On chemotherapy	9
Postchemotherapy	27
No chemotherapy	34
In relapse	
On chemotherapy	24
On erlotinib	31
No current treatment	33
Locally advanced	
On chemoradiation	14

TABLE 2. Mean EQ-5D Scores and QLQ-C30 Scores

	Mean ± SD	95% Bootstrap CI
EQ-5D		
Utility score	0.76 ± 0.20	0.73–0.78
VAS	67.22 ± 19.77	64.17–70.19
QLQ-C30		
Physical functioning	73.25 ± 21.81	70.08–76.54
Role functioning	67.44 ± 31.62	62.60–72.09
Emotional functioning	75.19 ± 24.05	71.51–78.68
Cognitive functioning	79.84 ± 22.67	76.36–83.14
Social functioning	73.16 ± 29.20	68.70–77.33
Global health status/QoL	65.89 ± 20.44	62.79–68.94
Fatigue	40.83 ± 25.94	37.02–44.70
Nausea and vomiting	7.56 ± 15.22	5.43–9.98
Pain	25.68 ± 28.80	21.51–29.84
Dyspnea	31.20 ± 30.19	26.74–36.05
Insomnia	34.88 ± 33.97	29.84–40.12
Appetite loss	22.67 ± 30.74	18.22–27.33
Constipation	18.99 ± 28.86	14.92–23.26
Diarrhea	13.37 ± 24.36	9.88–17.25
Financial difficulties	23.06 ± 33.30	18.41–28.10

VAS, visual analog scale; CI, confidence interval; QoL, quality of life.

TABLE 3. Coefficients for the Linear Regression Model

	Full Model	Reduced Model
Intercept	0.3381	0.4029
Physical functioning	0.0035*	0.0039*
Role functioning	0.0007	0.0008†
Emotional functioning	0.0011†	0.0015‡
Cognitive functioning	0.0007	
Social functioning	−0.0007	−0.0007
Global health status/QoL	0.0009	
Fatigue	0.0003	
Nausea and vomiting	−0.0002	
Pain	−0.0021*	−0.0021*
Dyspnoea	−0.0001	
Insomnia	−0.0001	
Appetite loss	−0.0001	
Constipation	0.0005	
Diarrhea	0.0004	
Financial difficulties	−0.0001	

* $p < 0.001$; † $p < 0$; ‡ $p < 0.01$.
QoL, quality of life.

majority had stage IV. Table 2 lists mean EORTC QLQ-C30 scores by functional and symptom scales.

The final linear model after stepwise variable reduction based on AIC included four functioning scales (physical, role, emotional, and social) and one symptom scale (pain). Physical functioning and pain had the largest absolute coefficient values, with values of 0.0039 and −0.0021, respectively, in the reduced model. Other coefficients for the original full model and reduced models are presented in Table 3. Adjusted R^2 values for the linear models were 0.57 and 0.58 for the full and reduced linear models, respectively.

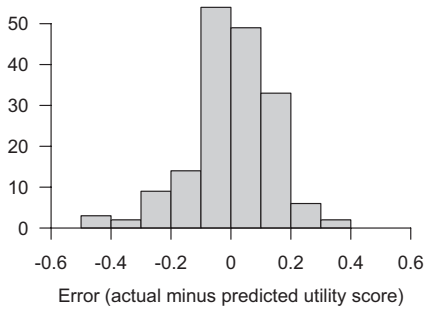

FIGURE 1. Histogram of the estimated errors (actual minus predicted utility scores).

The mean square error using 10-fold cross-validation was 0.02. Figure 1 shows the distribution of the estimated errors (actual minus predicted utility scores), whereas Figure 2 shows scatter plots with actual and predicted utility scores.

Mean predicted utility score based on the cross-validation was 0.76, which was the same as actual mean utility score. Subgroup analysis was performed by disease stage and disease state (Table 4), as well as the mean and 95% bootstrap CIs for the difference between the actual and predicted scores. The 95% bootstrap CI did not include 0, hence representing poor prediction for the subgroup in relapse and on erlotinib.

DISCUSSION

With the escalating cost of novel cancer therapies, cost-utility analyses have become increasingly more important. Currently, many clinical trials do not collect utility values but do collect QoL data through instruments such as the QLQ-C30 and FACT-G. Hence, there has been growing interest in models that reliably convert QoL data into utility scores. We studied the application of a linear regression model in patients with NSCLC to convert from the EORTC QLQ-C30 data to utility scores derived via the EQ-5D.

The linear regression model performed well and represents a promising approach of converting QoL data into utility scores. Of the 15 QLQ-C30 dimensions, 4 function dimensions (physical, role, emotional, and social) and the pain dimension were most predictive of utility score, as evidenced by their inclusion in the reduced model. Of these dimensions, physical function and pain had the most weight in the calculation of the utility score. The QLQ-C30 dimensions in the reduced model have significant overlap with the EQ-5D dimensions (e.g., the physical QLQ-C30 dimension corresponds to the mobility EQ-5D dimension and the emotional QLQ-C30 dimension corresponds to the anxiety/depression EQ-5D dimension), which is likely the reason why the dimensions in the reduced model best predict EQ-5D utility scores. However, by the same token, the QLQ-C30 dimensions in the reduced model cannot be considered to be the only important factors contributing to a patient's utility score, because there may be other dimensions such as a patient's financial difficulties that are not captured in the EQ-5D model.

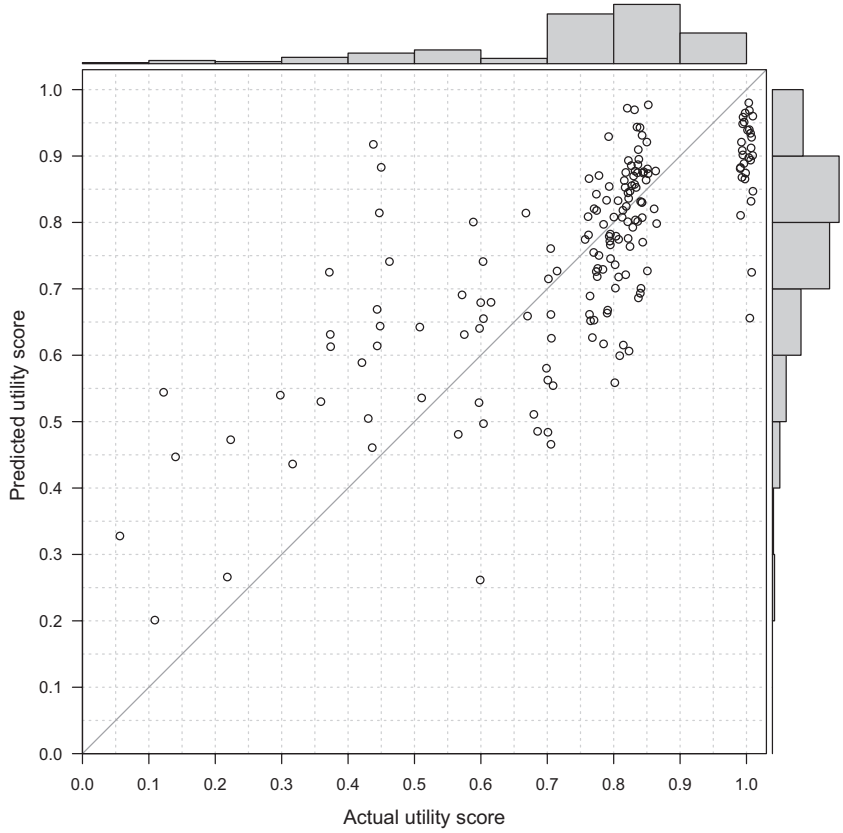

FIGURE 2. Scatter plot of actual versus predicted utility scores with corresponding histogram. The vertical distance between the point and the diagonal line represents the error between actual and predicted utility scores.

TABLE 4. Actual and Predicted Utility Scores, by Subgroup

	Actual Utility Score	Predicted, Linear Model
Whole sample	0.76 ± 0.20	0.76 ± 0.15
Range	0.05–1.00	0.20–0.97
Disease stage		
Stage I (n = 34)	0.80 ± 0.18	0.80 ± 0.14
Stage II (n = 16)	0.78 ± 0.23	0.80 ± 0.12
Stage III (n = 36)	0.73 ± 0.23	0.74 ± 0.13
Stage IV (n = 86)	0.75 ± 0.15	0.77 ± 0.13
Relapse free		
Chemotherapy (n = 9)	0.76 ± 0.04	0.74 ± 0.06
Post chemotherapy (n = 27)	0.76 ± 0.21	0.76 ± 0.12
No chemotherapy (n = 34)	0.77 ± 0.22	0.80 ± 0.16
Relapse		
Chemotherapy (n = 24)	0.69 ± 0.25	0.72 ± 0.18
Erlotinib (n = 31)	0.77 ± 0.17	0.73 ± 0.16
No current treatment (n = 33)	0.75 ± 0.20	0.75 ± 0.16
Locally advanced		
Chemotherapy (n = 14)	0.78 ± 0.17	0.77 ± 0.14

To the best of our knowledge, this is the first study in NSCLC that derives utility scores from EORTC QLQ-C30 data using the EQ-5D in patients. In patients with metastatic prostate cancer, a study mapping FACT-P and EORTC QLQ-C30 to EQ-5D utility scores obtained a R^2 value of 0.582 for its best model, meaning that it explained 58.2% of the observed variation in the validation sample.[13] Other articles that map EORTC QLQ-C30 to utilities using the EQ-5D have been recently published with comparable outcomes in breast,[14] gastric,[15] and esophageal cancers.[16] Our best model was developed through stepwise variable reduction to ensure a balance between model complexity and goodness-of-fit. Additional variables in the full model may overfit the data, compared with the reduced model. A small improvement in goodness-of-fit, assessed by R^2, was observed with our reduced model (0.58) in contrast to the full model (0.57).

Limitations of this study include a relatively small sample size and the lack of a unique population set to test for external validity. The population tariffs were also based on a subset of the general US population, which may not appropriately represent health preference in Canadian patients with NSCLC. The mean EQ-5D utility (0.76 ± 0.20) was surprisingly high, given that half of the patients had stage IV disease, and little variability was found across different disease states. The high utility score may reflect a biased sample of higher performance status patients who were willing to complete the questionnaires. However, similar scores to ours were obtained in one of the few studies to calculate EQ-5D utility scores in patients with lung cancer, in which the mean utility score was 0.74 ± 0.15 using the same US tariffs in patients with stage III or IV disease.[17] Using UK tariffs, the mean utility score was 0.67 ± 0.22. In that study, utility scores decreased with worsening performance status. However, in our study, we did not collect data on performance

status, and subgroup analysis by disease status may not have as large an impact on utility scores as performance status.

Finally, few patients enrolled in the study had reported low utility scores (e.g., <0.4). This may be reflected in the poorer performance of the models for patients who had low utility scores.

CONCLUSION

This study demonstrates the feasibility of converting QoL data into utilities in patients with NSCLC using linear modeling. If the predictive QLQ-C30 variables found in our study is validated in an independent series of patients, then EORTC QLQ-C30 data could be used to derive utility scores in patients with NSCLC. This may, in turn, further the ability to estimate cost utility of therapies for economic evaluations when utility values are not collected during the clinical trial.

REFERENCES

1. 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.
2. Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *J Clin Oncol* 1993;11:570–579.
3. Revicki DA. Relationship between pharmacoeconomics and health related quality of life. In B Spilker (Ed.), Quality of Life and Pharmacoeconomics in Clinical Trials. Philadelphia, PA: Lippincott-Raven Publishers, 1996. Pp. 1077–1083.
4. Torrance GW. Designing and conducting cost-utility analyses. In Spilker B (Ed.), Quality of Life and Pharmacoeconomics in Clinical Trials. Philadelphia, PA: Lippincott-Raven Publishers, 1996. Pp. 1105–1111.
5. Johnson FR, Hauber AB, Osoba D. Are chemotherapy patients' HRQoL importance weights consistent with linear scoring rules? A stated-choice approach. *Qual Life Res* 2006;15:285–298.
6. Mortimer D, Segal L. Comparing the incomparable? A systematic review of competing techniques for converting descriptive measures of health status into QALY-weights. *Medi Decis Making* 2008;28:66–89.
7. Bagust A, Barraza-Llorens M, Philips Z. Deriving a compound quality of life measure from the EORTC-QLQ-C30/LC13 instrument for use in economic evaluations of lung cancer clinical trials. *Eur J Cancer* 2001;37:1081–1088.
8. Fayers PM, Aaronson NK, Bjordal K. The EORTC QLQ-C30 Scoring Manual, 3rd Ed. Brussels: European Organisation for Research and Treatment of Cancer, 2001.
9. Kind P. The EuroQoL instrument: an index of health-related quality of life. In B Spilker (Ed.), Quality of Life and Pharmacoeconomics in Clinical Trials. Philadelphia, PA: Lippincott-Raven Publishers, 1996. Pp. 191–201.
10. EuroQoL Group. EuroQoL—a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
11. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D2 valuation model. *Med Care* 2005;43:203–220.
12. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. New York, NY: Springer, 2001.
13. Wu EQ, Mulani P, Farrell MH, et al. Mapping FACT-P and EORTC QLQ-C30 to patient health status measured by EQ-5D in metastatic hormone-refractory prostate cancer patients. *Value Health* 2007;10:408–414.
14. Crott R, Briggs A. Mapping the QLQ-C30 quality of life cancer questionnaire to EQ-5D patient preferences. *Eur J Health Econ* 2010;11:427–434.
15. Kontodimopoulos N, Aletras VH, Paliouras D, et al. Mapping the cancer-specific EORTC QLQ-C30 to the preference-based EQ-5D, SF-6D, and 15D Instruments. *Value Health* 2009;12:1151–1157.
16. McKenzie L, van der Pol M. Mapping the EORTC QLQ C-30 onto the EQ-5D instrument: the potential to estimate QALYs without generic preference data. *Value Health* 2009;12:167–171.
17. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health Qual Life Outcomes* 2007;5:70.

