Mapping the cancer-specific EORTC QLQ-C30 and EORTC QLQ-BR23 to the generic EQ-5D in metastatic breast cancer patients

Eun-ju Kim · Su-Kyoung Ko · Hye-Young Kang

Abstract

Purpose To develop a mapping algorithm for a conversion of the EORTC QLQ-C30 and EORTC QLQ BR-23 into the EQ-5D-derived utilities in metastatic breast cancer (MBC) patients.

Methods We enrolled 199 patients with MBC from four leading Korean hospitals in 2009. EQ-5D utility, cancer-specific (QLQ-C30) and breast cancer-specific quality of life data (QLQ-BR23) and selected clinical and demographic information were collected from the study participants. Ordinary least squares regression models were used to model the EQ-5D using QLQ-C30 and QLQ-BR23 scale scores. To select the best model specification, six different sets of explanatory variables were compared.

Result Regression analysis with the multiitem scale scores of QLQ-C30 was the best-performing model, explaining for 48.7% of the observed EQ-5D variation. Its mean absolute error between the observed and predicted EQ-5D utilities (0.092) and relative prediction error (2.784%) was among the smallest. Also, this mapping model showed the least systematic errors according to disease severity.

Conclusions The mapping algorithms developed have good predictive validity, and therefore, they enable researchers to translate cancer-specific health-related quality of life measures to the preference-adjusted health status of MBC patients.

Keywords EORTC QLQ-C30 · EORTC QLQ-BR 23 · EQ-5D · Mapping · Utility · Quality of life

Abbreviations

CUA	Cost-utility analysis.
ECOG	Eastern cooperative oncology group
EORTC QLQ-BR23 (or QLQ-BR23)	European Organization for Research and Treatment of Cancer quality of life questionnaire breast cancer-23
EORTC QLQ-C30 (or QLQ-C 30)	European Organization for Research and Treatment of Cancer quality of life questionnaire core-30
EQ-VAS	EQ-5D visual analogue scale
ISPOR	International society for pharmacoeconomics and outcomes research
MBC	Metastatic breast cancer
NHI	National health insurance
OLS	Ordinary least squares
QALYs	Quality-adjusted life years
QOL	Quality of life
RPE	Relative prediction error
VIF	Variance inflation factor

E. Kim
Graduate School of Public Health, Seoul National University, Seoul, South Korea

S.-K. Ko
Graduate School of Public Health, Yonsei University, Seoul, South Korea

S.-K. Ko
Department of Market Access, Pfizer Pharmaceuticals Korea Limited, Seoul, South Korea

H.-Y. Kang (✉)
College of Pharmacy, Yonsei Institute of Pharmaceutical Sciences, Yonsei University, 162-1 Songdo-dong, Yeonsu-gu, Incheon 406-840, South Korea
e-mail: hykang2@yonsei.ac.kr; hykang@yuhs.ac

Introduction

Cost-utility analysis (CUA) is often used to assess treatment alternatives. The outcomes of CUA are measured in terms of quality-adjusted life years (QALYs), which are calculated on the basis of a single preference-based score. When measuring patients' quality of life (QOL), disease-specific instruments are often preferred to generic ones because they focus on particular health problems and they tend to be more responsive to clinically meaningful differences [1, 2]. While generic measures such as EQ-5D and SF-6D are preference-based measures of health and provide interval-scaled utility information, where 1 and 0 refer to full health and death, respectively, most of the disease-specific measures are not preference-based measures and do not have utility scoring systems. However, CUA requires a single preference-based score so that QALYs can be calculated. One solution to this problem is to attempt to predict the utilities by "mapping," i.e., regressing preference-based indices against the dimension or item scores of disease-specific measures, and thus to obtain estimation models that can be used to subsequently calculate QALYs [3, 4].

The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30) is the most commonly used cancer-specific quality of life questionnaire, and it has been cross-culturally validated [5, 6]. Mapping algorithms predicting the EQ-5D from the scale scores of the EORTC QLQ-C30 have been developed for various types of cancer, including gastric [1], esophageal [7], prostate [8] and breast cancer [9].

Breast cancer is the third most prevalent cancer in South Korea, and its prevalence increased sharply between 2003 and 2007 from 122.3 to 155.8 per 100,000 women [10], resulting in a substantially increased financial burden on Korean society. Subsequently, the choice of effective and efficient treatment alternatives to treat patients with breast cancer is a critical concern from the perspective of the Korean National Health Insurance (NHI) program. In particular, after the introduction of the "Positive Drug Listing System" in Korea in 2007, the NHI requires drug companies to submit cost-effectiveness evidence for newly approved pharmaceuticals in the Korean market to aid in determining their inclusion on the covered medicine list. As recommended by the Korean Health Insurance Review and Assessment Services [11], cost-utility analysis is a standard approach to evaluate the economic value of the treatment alternatives in Korea. However, the lack of information on utility values for the disease of interest hinders the ability to provide adequate evidence of cost-utility. Thus, developing a mapping algorithm to predict utility, such as EQ-5D, based on a cancer-specific QOL questionnaire among breast cancer patients, is in critical need in order to carry out CUA for comparing various

treatment options in Korea. Although a mapping algorithm for breast cancer patients was developed in an earlier study [9], cultural differences may limit the extent to which the mapping equations developed for European populations can be applied in Korea. Furthermore, the previous algorithm was estimated using a general cancer-specific quality of life questionnaire (i.e., EORTC QLQ-C30) only, but not with incorporating a breast cancer-specific quality of life questionnaire (i.e., EORTC QLQ-BR23), which is considered to be more sensitive and responsive to the functions and symptoms associated with breast cancer.

Therefore, this study was carried out to develop a mapping algorithm to transform the cancer-specific and breast cancer-specific QOL questionnaire responses into EQ-5D-derived utilities for breast cancer patients in Korea. In order to assess the health-related QOL of breast cancer patients, we used a cancer-specific QOL instrument (i.e., EORTC QLQ-C30) as well as a breast cancer-specific instrument (i.e., EORTC QLQ-BR23). In addition to the EORTC QLQ-C30, EORTC QLQ supplementary modules have been developed for selected cancers, including breast cancer, head and neck cancer and lung cancer, to assess the following: (1) symptoms related to a specific tumor site; (2) side effects associated with a given treatment; and (3) additional quality of life domains affected by the disease or treatment [12]. Regression analysis was used to establish the relationship between the cancer-specific QOL measures and the preference-based utility measures. The performance of the mapping model assessed how well the responses to the EORTC QLQ-C30 and EORTC QLQ-BR23 predict the EQ-5D responses for a separate data set.

Methods

Instruments

The EORTC QLQ-C30 (hereafter, QLQ-C30) is a popular instrument for measuring the general cancer quality of life [6], and it is used internationally [5]. The QLQ-C30 includes 30 items and is composed of 15 multiitem scales or quality of life domains that evaluate functioning, symptoms and overall health (2 items). The functioning scales include physical (5 items), emotional (4 items), cognitive (2 items), role (2 items) and social functioning (2 items) [12]. The symptom scales measure nausea and vomiting (2 items), fatigue (3 items) and pain (2 items) and 6 single items assessing financial impact and various physical symptoms. Most of the 30 items have 4 response levels (not at all, a little, quite a bit and very much), with 2 questions that address overall health with 7 response levels (on a scale from very poor to excellent). All of the scales

and single-item measures were scored according to the standard scoring rules identified in the EORTC QLQ-C30 Scoring Manual [12]. The principles for scoring these scales are the same in all cases. First, we estimate the average of the items that contribute to the scale; this is the raw score. Second, we use a linear transformation to standardize the raw score, so that the scores range from 0 to 100 (for the technical summary, see "Appendix 1"). Higher functional scale scores represent better functioning, but higher symptom scores indicate worse symptoms.

The EORTC QLQ-BR23 (hereafter, QLQ-BR23) is a breast cancer module of the EORTC QLQ and is developed for use among breast cancer patients varying in disease stage and treatment modality (i.e., surgery, chemotherapy, radiotherapy and endocrine treatment). When employed in conjunction with the QLQ-C30, the use of the QLQ supplementary modules, including QLQ-BR23, can provide more detailed information relevant to evaluating the QOL in specific patient populations. It includes 23 items composed of 4 functioning scales [i.e., body image (4 items), sexual functioning (2 items), sexual enjoyment (1 item) and future perspective (1 item)] and 4 symptom scales [systemic therapy side effects (7 items), breast symptoms (4 items), arm symptoms (3 items) and upset by hair loss (1 item)] [13]. The scoring approach is identical in principle to that for the function and symptom scales and the single items of the QLQ-C 30 [13] (see "Appendix 1" for the technical summary).

EQ-5D is a commonly used generic instrument, which contains a 5-item classifier (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), plus a visual analogue scale (EQ-VAS) of the overall health status. Each classification item has 3 functional levels: no problem, some problems and extreme problems [14].

Sample and data collection

The study was a cross-sectional survey conducted in 4 leading hospitals in treating breast cancer in Korea, all of which were located in Seoul, the capital city of Korea: Asan Medical Center, National Cancer Center Hospital, Seoul National University Hospital and Yonsei Cancer Center. On a prospective basis, we enrolled patients with metastatic breast cancer (MBC), who were receiving palliative chemotherapy from the participating hospitals between May and August 2009. The study includes 199 patients who provided informed consent and completed all of the 3 questionnaires and common demographic questions during their hospital visits. Disease-related data were taken from the patients' medical chart. All of the surveys were approved by the Institutional Review Boards from the participating hospitals.

Using an interviewer-assisted self-administered questionnaire survey mode, each patient completed the Korean version of the QLQ-C30 and QLQ-BR23 questionnaires to assess their QOL, and they also answered the Korean version of the EQ-5D questionnaire to assess their preference or utility for their current health status with MBC. The patients also retrospectively completed the QLQ-C30 and QLQ-BR23 questionnaires for their health status before they were diagnosed with breast cancer. For this study, we purchased the official Korean version of the QLQ-C30 and QLQ-BR23 from the European Organization for Research and Treatment of Cancer (EORTC, www.eortc.org) and we also purchased the official Korean version of the EQ-5D from the EuroQol group (http://www.euroqol.org/eq-5d/eq-5d-products/eq-5d-3l-translations.html). Utility ranged from 1, "perfect health" to 0, "death" to negative values representing health states the person considers worse than death. We used the Korean EQ-5D index scoring of Lee et al. [15] to calculate a single summary score from 5 descriptive items of the EQ-5D.

Data analysis

Ordinary least squares (OLS) regression was used to model the EQ-5D using the QLQ-C30 and QLQ-BR23 scale scores and selected patient demographic and clinical characteristics as predicting variables. We checked for deviations from the main OLS regression assumptions. The responses for the current health status of the patients with MBC were used for the model estimation. The EQ-5D summary score was the dependent variable of the models. The models defined in this study are additive, implying linear independence between predictor variables. We pursued a simple-to-use and robust algorithm that utilized a few demographic and clinical variables, which are easily accessible and accurate in routine clinical practice including age, gender and Eastern Cooperative Oncology Group (ECOG) performance score. ECOG performance score was incorporated into the regression model to reflect the baseline severity of the patient's health condition [16, 17]. Predicted values of EQ-5D exceeding 1.0 were truncated to this boundary value. To select the best model specification, six different sets of explanatory variables were compared:

Model 1: QLQ-C 30 scale scores (15 scales)
Model 2: QLQ-C 30 scale scores (15 scales) and demographics (age, gender and ECOG score)
Model 3: QLQ-BR23 scale scores (6 out of 8 scales after excluding 2 scales of "sexual enjoyment" and "upset by hair loss")
Model 4: QLQ-BR23 scale scores (6 scales) and demographics (age, gender and ECOG score)
Model 5: QLQ-C 30 scale scores (15 scales) and QLQ-BR 23 scale scores (6 scales)

Model 6: QLQ-C 30 scale scores (15 scales), QLQ-BR23 scale scores (6 scales) and demographics (age, gender and ECOG score)

While there was a 100% response rate for the QLQ-C30 question items, the study participants did not do the same for the QLQ-BR23 items. For the two questions of the QLQ-BR23, which are "sexual enjoyment" and "upset by hair loss," only 41 and 145 patients, respectively, answered in regard to their current health status with MBC. The low response rate for these two questions was due to the following targeted questions: "Answer this question only if you had any hair loss. Were you upset by the loss of your hair?" and "Answer the question only if you have been sexually active. To what extent was sex enjoyable for you?" Thus, the patients who did not answer the hair loss question could have omitted this because they had not lost hair. Similarly, those not answering the sexual enjoyment question may not be sexually active. If we estimate the regression models based on the observations with complete responses for all of the variables, then the loss of sample size is considerable and the statistical power of the estimation model would not be stable. Therefore, we excluded these two scales of QLQ-BR 23 from the model and estimated the Models 3, 4, 5 and 6.

Selection of the best models was based on the predictive performance of the models. Because a prediction model usually performs better with the data that were used in its development, it is recommended to evaluate how well the model predicts the EQ-5D responses with other data sets [8, 18]. Thus, the total sample of 199 respondents in this study was split into two groups using computer-generated random numbers: 75% of the patients (i.e., 149 patients) were randomly assigned to the "estimation sample" that was used to generate the mapping models. The remaining 25% (i.e., 50 patients) were used as the "validation sample" to test the performance of the models produced. Various approaches were used to test the performances of the models. First, the absolute deviation between the actual and predicted EQ-5D summary scores, or the mean absolute error, served as the primary measurement of the model's performance. Second, many mapping algorithms have been found to have systematic errors in prediction that they usually predict higher values than the actual EQ-5D values for patients who are in poor health, but lower values than the actual values for patients with mild health problems [19–21]. In order to examine how the prediction accuracy varies with the disease severity, we split the validation sample into two groups: (1) the subjects with worse health states, which is defined as having EQ-5D utility score lower than the average (0.799) of the validation sample, and (2) the subjects with better health statues that have EQ-5D utility score higher than the average.

Then, the mapping models with the least systematic errors, i.e., the least mean absolute errors for both groups with worse or better health states, were selected as the best models. Last, the relative prediction error (RPE), which is calculated by dividing the difference between actual and predicted utility values by the actual value, was taken as the performance criterion of the models, with the smaller the value, the better the performance [22].

Results

Descriptive statistics for the selected characteristics of the study participants are presented in Table 1. The mean age

Table 1 Baseline characteristics of the study participants

Characteristics	N (%)
Age (years)	
Mean (range)	49.3 (26–77)
<50	100 (50.3)
≥50	99 (49.7)
Gender	
Female	196 (98.5)
Male	3 (1.5)
Menopausal status	
Pre-menopausal	45 (22.6)
Post-menopausal	140 (70.4)
Missing/no response	14 (7.0)
ECOG performance status, grade	
0	44 (22.1)
1	142 (71.4)
2	11 (5.5)
3	0 (0)
4	1 (0.5)
Missing/no response	1 (0.5)
Best response to the latest regimen	
Progressive disease	10 (5.0)
CR, PR, SD	168 (83.8)
Not evaluable	21 (11.2)
Duration of metastasis: mean (range) years	2.9 (0–16.2)
Number of metastases	
Mean (SD)	1.7 (0.8)
1	98 (49.3)
2	68 (34.2)
3	27 (13.6)
4	4 (1.9)
Missing/no response	2 (1.0)

ECOG eastern cooperative oncology group, *CR* complete response, *PR* partial response, *SD* stable disease

of the study participants was 49.3 years old. Three patients were male, and 70.4% of the study participants were postmenopausal. The majority of the patients fell in the ECOG performance grade of 0 (22.1%) or 1 (71.4%). On average, the patients spent 2.9 years in a state of metastasis with a mean number of metastases equivalent to 1.7.

The mean utility score derived by EQ-5D among 199 study subjects was 0.82 (SD = 0.16) for the current health state with MBC, reduced by about 15% from the utility score of 0.96 (SD = 0.16) for the pre-cancer health state (Table 2). The global health status, which is a component of the QLQ-C30, was decreased by 29% from 74.58 to 53.27 on a 100 scale after diagnosing breast cancer. According to the QLQ-C30 measurement, breast cancer appears to bring more deterioration in symptoms than in functioning. While the quality of life score for physical, emotional, cognitive, role and social functioning was reduced by 12 to 34%, the scores for the symptoms were increased or worsened by a bigger amount ranging from 52 to 503%. A similar trend was observed in the QLQ-BR23 measurement: There was a 20–67% reduction in functioning and a 72–278% worsening in symptoms.

Using the regression coefficients estimated by each model, the predicted EQ-5D utility scores were calculated for the individual patients of the validation sample. The mean predicted EQ-5D values from each model ranged from 0.598 to 0.824, compared with 0.799 for the observed EQ-5D value of the validation sample (Table 3). To select the best model in terms of the predictive performance, the mean absolute error was used as the primary indicator. Also, the relative prediction error was compared across all the prediction models (Table 3). The best-performing model is Model 1, which is an OLS model using scale scores of the QLQ-C 30 as predictors. This model's mean absolute error (MAE) (0.092) is smallest among all of the estimates. Model 1 has the least systematic error for prediction according to the disease severity. Also, the relative prediction error (RPE, 2.784%) was among the smallest. Finally, the goodness of fit of Model 1, as measured by the adjusted R^2 (0.487), was reasonably high (Table 4).

Although the mean absolute error of the Models 5 and 6, which included both of the QLQ-C30 and QLQ-BR 23 scale scores, is similar to the Model 1, all of the QLQ-BR 23 scale scores and majority of the QLQ-C30 scale scores were dropped from the model due to multicollinearity problem. Variables with the variance inflation factor (VIF) of 10 or higher were dropped from all models, which is a criterion used to determine multicollinearity in our analysis [23, 24]. The explanatory power and fit of the Models 3 and 4 was relatively low with adjusted R^2 of 0.246 and 0.235, respectively.

Table 5 shows the regression analysis results of each mapping model. In this section, our interpretation of the regression coefficients focuses on the best-performing

model, which is Model 1. Three scales of the QLQ-C30 are statistically significant ($P < 0.05$): physical functioning ($\beta = 0.0032$), pain ($\beta = -0.0013$) and dyspnea ($\beta = 0.0008$). Some of the scales were not retained in the model due to multicollinearity problem: role functioning, cognitive functioning, fatigue symptom, nausea/vomiting, insomnia, constipation and financial difficulties. "Appendix 2" provides covariance matrix for the best-performing model, Model 1.

Discussion

The purpose of this study was to develop mapping algorithms to be applied to patients with MBC in order to calculate the utilities at various time points during the course of a treatment, and thus, these algorithms can be used to derive the QALYs. Although the prediction models based on the QLQ-C30 have been estimated in earlier studies, those were for other cancer patients such as those with gastric [1], esophageal [7] and prostate cancer [8]. The present study is the first, to the best of our knowledge, to attempt to predict the EQ-5D utility index based on both QLQ-C30 and QLQ-BR 23 scales of breast cancer patients. There is a recent study that developed a mapping algorithm for patients with breast cancer, but it used only the QLQ-C30 and not the QLQ-BR23 [9]. As compared to the mapping algorithm developed for patients with MBC based on a European population [9], the mapping equations developed from this study have an advantage of improving their applicability to the Korean population.

The best-performing model (Model 1) developed in the present study had good predictive validity for patients with breast cancer. The mean absolute deviation between the observed and predicted EQ-5D utility scores of the best-performing model in this study was the smallest among the proposed 6 prediction models. Also, the RPE of the Model 1 (2.784%) was comparable to that of earlier studies (4.649 and 3.787%) [1, 7]. In the process of selecting models, we found overall that the models with the QLQ-C30 only had a better predictive performance and contain more scales in the model than those with both of the QLQ-C30 and QLQ-BR23. For example, the models that included the QLQ-BR 23 only (Model 3) or the QLQ-BR 23 combined with selected demographic and clinical variables (Model 4) as predictors had a higher mean absolute error (0.109 and 0.109) than did the models that included the QLQ-C30 only (Model 1) or the QLQ-C30 combined with selected demographic and clinical variables (Model 2) (0.092 and 0.095). Also, as seen in the regression analysis results in Table 5, all of the QLQ-BR23 scales and about 75% of the QLQ-C30 scales were dropped from the model due to high correlation between the scales causing multicollinearity

Table 2 Descriptive statistics for health utility and the health-related quality of life variables

Variable	Health status before the diagnosis of BC		Health status with MBC		Change (%)
	N (%)	Mean (SD)	N (%)	Mean (SD)	
EQ-5D		0.96 (0.06)		0.82 (0.16)	−15
Mobility	199		199		
No problems	185 (92.96)		102 (51.26)		
Moderate problems	13 (6.53)		92 (46.23)		
Severe problems	1 (0.51)		5 (2.51)		
Self-care	199		199		
No problems	199 (100.0)		158 (79.40)		
Moderate problems			38 (19.10)		
Severe problems			3 (1.50)		
Usual activities	199		199		
No problems	187 (93.97)		72 (36.18)		
Moderate problems	12 (6.03)		123 (61.81)		
Severe problems			4 (2.01)		
Pain	199		199		
No problems	156 (78.39)		47 (23.62)		
Moderate problems	40 (20.10)		137 (68.84)		
Severe problems	3 (1.51)		15 (7.54)		
Anxiety/depression	199		199		
No problems	130 (65.33)		60 (30.15)		
Moderate problems	64 (32.16)		127 (63.82)		
Severe problems	5 (2.51)		12 (6.03)		
EORTC QLQ-C30					
Global health status	199	74.58 (21.33)	199	53.27 (22.23)	−29
Physical functioning	199	92.83 (11.05)	199	65.29 (20.04)	−30
Role functioning	199	89.78 (18.55)	199	59.21 (28.31)	−34
Emotional functioning	199	75.63 (24.83)	199	66.29 (23.62)	−12
Cognitive functioning	199	89.36 (16.66)	199	69.68 (21.43)	−22
Social functioning	199	88.86 (21.40)	199	58.21 (28.91)	−34
Fatigue symptom	199	25.20 (22.4)	199	48.46 (24.21)	92
Nausea/vomiting	199	3.18 (11.28)	199	19.18 (23.38)	503
Pain	199	10.80 (19.80)	199	34.51 (26.62)	220
Dyspnea	199	8.21 (16.57)	199	28.64 (29.02)	249
Insomnia	199	16.42 (25.49)	199	33.84 (31.87)	106
Appetite loss	199	7.20 (17.36)	199	34.51 (32.37)	379
Constipation	199	15.91 (24.58)	199	24.12 (29.38)	52
Diarrhea	199	7.71 (17.63)	199	16.25 (22.93)	111
Financial difficulties	199	12.40 (26.64)	199	47.24 (36.57)	281
EORTC QLQ-BR23					
Body image	199	91.27 (16.33)	199	46.40 (29.92)	−49
Sexual functioning	199	28.14 (25.15)	199	9.30 (17.04)	−67
Sexual enjoyment	122	42.62 (17.28)	41	34.15 (17.46)	−20
Future perspective	199	80.74 (26.44)	199	34.84 (32.53)	−57
Systemic therapy side effects	199	13.42 (12.79)	199	36.49 (18.14)	172
Breast symptoms	199	10.18 (16.26)	199	17.84 (16.90)	75
Arm symptoms	199	9.38 (14.52)	199	35.46 (25.74)	278
Upset by hair loss	40	41.67 (30.89)	145	71.49 (34.24)	72

BC breast cancer, *MBC* metastatic breast cancer, *EORTC QLQ-C30* European Organization for Research and Treatment of Cancer quality of life questionnaire core-30, *EORTC QLQ-BR23* European Organization for Research and Treatment of Cancer quality of life questionnaire breast cancer-23

Table 3 Performance of the models to predict the EQ-5D values

Specification of the models	Mean actual EQ-5D value (SD)	Mean predicted EQ-5D value (SD)	Mean absolute error (SD)	Mean relative prediction error %
Model 1: EORTC QLQ-C 30	0.799 (0.160)	0.821 (0.117)	0.092 (0.082)	2.784
Model 2: EORTC QLQ-C 30 + demographics	0.799 (0.160)	0.824 (0.113)	0.095 (0.088)	3.203
Model 3: EORTC QLQ-BR23	0.799 (0.160)	0.823 (0.082)	0.109 (0.117)	3.017
Model 4: EORTC QLQ-BR23 + demographics	0.799 (0.160)	0.819 (0.082)	0.109 (0.117)	2.557
Model 5: EORTC QLQ-C 30 + QLQ-BR 23	0.799 (0.160)	0.815 (0.113)	0.094 (0.080)	1.988
Model 6: EORTC QLQ-C 30 + EORTC QLQ-BR 23 + demographics	0.799 (0.160)	0.598 (0.169)	0.109 (0.117)	−25.157

$N = 50$ (validation sample)

Mean absolute error = Mean of absolute value of (predicted utility for patient "n"-actual utility for patient "n")

Mean relative predication error % = Mean of [{(predicted utility for patient "n"-actual utility for patient "n")/actual utility for patient "n"} × 100]

EORTC QLQ-C30 European Organization for Research and Treatment of Cancer quality of life questionnaire core-30, *EORTC QLQ-BR23* European Organization for Research and Treatment of Cancer quality of life questionnaire breast cancer-23

Table 4 Performance of the models to predict the EQ-5D values according to the disease severity

Specification of the model	Mean actual EQ-5D value (SD)	Mean predicted EQ-5D value (SD)	Mean absolute error (SD)	Mean relative prediction error %
Patients in the validation sample with worse health states [EQ-5D score < 0.799 (=mean score of the validation sample)]: $n = 22$				
Model 1: EORTC QLQ-C 30	0.665 (0.147)	0.755 (0.110)	0.119 (0.100)	13.489
Model 2: EORTC QLQ-C 30 + demographics	0.665 (0.147)	0.761 (0.099)	0.126 (0.107)	14.329
Model 3: EORTC QLQ-BR23	0.665 (0.147)	0.792 ((0.076)	0.149 (0.160)	19.060
Model 4: EORTC QLQ-BR23 + demographics	0.665 (0.147)	0.781 (0.072)	0.149 (0.160)	17.410
Model 5: EORTC QLQ-C 30 + QLQ-BR 23	0.665 (0.147)	0.753 (0.106)	0.118 (0.099)	13.231
Model 6: EORTC QLQ-C 30 + EORTC QLQ-BR 23 + demographics	0.665 (0.147)	0.753 (0.106)	0.118 (0.099)	13.231
Patients in the validation sample with better health states [EQ-5D score ≥ 0.799 (=mean score of the validation sample)]: $n = 28$				
Model 1: EORTC QLQ-C 30	0.901 (0.063)	0.873 (0.095)	0.072 (0.059)	−3.411
Model 2: EORTC QLQ-C 30 + demographics	0.901 (0.063)	0.874 (0.098)	0.070 (0.061)	−3.236
Model 3: EORTC QLQ-BR23	0.901 (0.063)	0.847 (0.079)	0.078 (0.053)	−6.266
Model 4: EORTC QLQ-BR23 + demographics	0.901 (0.063)	0.849 (0.078)	0.078 (0.053)	−6.038
Model 5: EORTC QLQ-C 30 + QLQ-BR 23	0.901 (0.063)	0.863 (0.095)	0.075 (0.058)	−4.517
Model 6: EORTC QLQ-C 30 + EORTC QLQ-BR 23 + demographics	0.901 (0.063)	0.863 (0.095)	0.075 (0.058)	−4.517

Mean absolute error = Mean of absolute value of (predicted utility for patient "n"-actual utility for patient "n")

Mean relative predication error % = Mean of [{(predicted utility for patient "n"-actual utility for patient "n")/actual utility for patient "n"} × 100]

EORTC QLQ-C30 European Organization for Research and Treatment of Cancer quality of life questionnaire core-30, *EORTC QLQ-BR23* European Organization for Research and Treatment of Cancer quality of life questionnaire breast cancer-23

problems. Thus, unlike our initial hypothesis, we would conclude that general cancer-specific quality of life questionnaire (i.e., QLQ-C30) alone is more suitable to predict the EQ-5D utility than both of the general (i.e., QLQ-C30) and the breast cancer-specific quality of life questionnaire (i.e., QLQ-BR23) together.

Not all scales of the QLQ-BR23 in the best-performing model in this study have intuitive signs for their regression coefficients. For example, the positive coefficient for the dyspnea scale from Model 1 indicates a better quality of life if this symptom gets worse. Such counterintuitive signs may threaten the model reliability. However, from earlier studies that have developed mapping algorithms to predict the EQ-5D utilities using QLQ-C30 instruments, we have found the similar problems. None of the studies appeared to develop mapping models with perfectly intuitive signs for the regression coefficients of all the QLQ-C30 scales. For instance, from the study by Crott and Briggs (2010)

Table 5 Regression analysis results of the best mapping models for the EQ-5D values

	Regression coefficient (standard error)					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
EORTC QLQ-C30						
Intercept	0.5534 * (0.0614)	0.6043 * (0.01015)	0.9400 * (0.1045)	0.8980 * (0.0380)	0.5849 * (0.0579)	0.6340 * (0.1003)
Global health status (GH)	0.0007 (0.0005)	0.0004 (0.0005)			0.0010 (0.0005)	0.0008 (0.0005)
Physical functioning (PF)	0.0032 * (0.0006)	0.0033 * (0.0006)			0.0035 * (0.0006)	0.0037 * (0.0006)
Role function (RF)	NR	NR			NR	NR
Emotional functioning (EF)	0.0005 (0.0005)	0.0007 (0.0005)			NR	NR
Cognitive functioning (CF)	NR	NR			NR	NR
Social functioning (SLF)	0.0005 (0.0004)	0.0005 (0.0004)			NR	NR
Fatigue symptom (FS)	NR	NR			NR	NR
Nausea/vomiting (NV)	NR	NR			NR	NR
Pain (P)	−0.0013 * (0.0005)	−0.0014 * (0.0005)			−0.0014 * (0.0005)	−0.0015 * (0.0005)
Dyspnea (DS)	0.0008 * (0.0003)	0.0009 (0.0004)			0.0008 * (0.0003)	0.0008 (0.0004)
Insomnia (I)	NR	NR			NR	NR
Appetite loss (A)	−0.0006 (0.0003)	−0.0006 (0.0003)			−0.0006 (0.0003)	−0.0006 (0.0003)
Constipation (C)	NR	NR			NR	NR
Diarrhea (D)	0.0005 (0.0004)	0.0003 (0.0005)			NR	NR
Financial difficulties (F)	NR	NR			NR	NR
EORTC QLQ-BR23						
Body image (BI)			NR	NR	NR	NR
Sexual functioning (SF)			0.0020 * (00006)	0.0018 * (0.0007)	NR	NR
Sexual enjoyment (SE)			NR	NR	NR	NR
Future perspective (FP)			0.0007 (0.0004)	0.0008 * (0.0004)	NR	NR
Systemic side effects (SE)			−0.0015 * (0.0007)	−0.0014 * (0.0007)	NR	NR
Breast symptoms (BS)			NR	NR	NR	NR
Arm symptoms (AS)			−0.0017 * (0.0005)	−0.0017 * (0.0005)	NR	NR
Age: <50 (ref)						
≥50 years old	−0.0335 (0.0195)	−0.0335 (0.0195)		−0.0192 (0.0226)		−0.0308 (0.0185)
Gender: female (ref)						
Male	−0.0336 (0.0673)	−0.0336 (0.0673)		0.0289 (0.0815)		−0.0336 (0.0672)

Table 5 continued

| | Regression coefficient (standard error) | | | | | |
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
ECOG performance status						
Grades 0–1 (ref)						
Grades 2–4		0.0299 (0.0463)		−0.0494 (0.0555)		0.0366 (0.0464)
Adjusted R^2	0.4872	0.4874	0.2461	0.2351	0.4827	0.4813

$N = 149$ (estimation sample). *Statistically significant at 5% level

EORTC QLQ-C30 European Organization for Research and Treatment of Cancer quality of life questionnaire core-30, *EORTC QLQ-BR23* European Organization for Research and Treatment of Cancer quality of life questionnaire breast cancer-23, *NR* not retained in the regression analysis

Using the regression coefficients of the models, the predicted EQ-5D utility scores can be calculated for each individual as follows: For model 1, EQ-5D utility index = $0.5534 + 0.0007 \times GH + 0.0032 \times PF + 0.0005 \times EF + 0.0005 \times SLF − 0.0013 \times P + 0.0008 \times DS − 0.0006 \times A + 0.0005 \times D$

(Mapping the QLQ-C30 to the EQ-5D preferences for breast cancer patients), the regression coefficient for the "dyspnea" scale was negative, which represents that a high dyspnea scale (i.e., worse symptom) decreases the disutility (i.e., increases the utility or the quality of life) [9]. A mapping algorithm by McKenzie and colleagues (2009) to predict the EQ-5D preferences for esophageal cancer patients using the QLQ-C30 showed that the regression coefficient for selected symptom scales, including dyspnea, insomnia, appetite loss and constipation, is positive [7]. This indicates a better quality of life if the symptoms of these domains get worse. Another example can be found from the study by Wu et al. [8], which mapped the FACT-P and QLQ-C30 to the EQ-5D preferences for prostate cancer patients. The regression coefficient for the role function scale has a negative sign, implying a worse quality of life if the role functioning gets improved. Also, some of the symptom scales such as dyspnea, insomnia, appetite loss and diarrhea and the financial difficulty scale had a positive coefficient, which indicates a better quality of life if the symptoms of these domains get worse [8]. Such counter-intuitive signs for the regression coefficients would be due to unknown interrelationships between the scales of QLQ-C30 when we put them together in a regression model. It is interesting to observe that overall, the functioning and global health status scales have appropriate signs for their regression coefficients in most of the studies. On the other hand, there is a tendency that the symptom scales show counterintuitive signs for their regression coefficients. In particular, we could observe that the dyspnea, insomnia, appetite loss and diarrhea (or constipation in some studies) scales have common counterintuitive signs in the previous studies including our study.

The present study has several limitations. First, the study results may have limited generalizability because patients from only selected hospitals were included in the study. To improve the generalizability of the equation developed in this study, future studies with other samples would be helpful. Second, one drawback of this study is the small sample size. Out of 199 patients, 149 were used in the estimation sample. The predictive performance of the estimated models was confirmed with only 50 patients of the validation sample. The QOL questionnaire survey in this study was carried out along with other questionnaires that asked breast cancer patients about their willingness-to-pay for anticancer treatments, which requires an expensive face-to-face interview method because of difficulty in responding to questions. Therefore, it was not possible to consider a mailed questionnaire survey including more centers to increase the sample size for this study. Last, the use of the patients' pre-cancer health state to examine the magnitude of the deterioration of QOL due to breast cancer may have introduced a recall bias problem because some patients had been

diagnosed several years prior to the interview and they may not have been able to accurately recall their health status at that time. According to the 2007 report of the Korea National Health and Nutrition Examination Survey, the mean EQ-5D utility value of Korean women aged 50–59 years old is 0.909 ($\pm$0.007) [25]. Using this value as a reference health state, the reduction in QOL due to breast cancer (9.8%) is smaller than that calculated using the mean EQ-5D utility value of the study patients for their pre-cancer state (15%). Therefore, the magnitude of the deterioration of QOL due to breast cancer reported in this study may have been overestimated.

Conclusions

The models we developed have good predictive validity. These algorithms enable researchers to translate cancer-specific health-related quality of life measures to a preference-adjusted health status of MBC patients. The findings will help perform cost-utility analyses of medical interventions for MBC in Korea.

Appendix 1: Technical summary: principles of scoring of QLQ-C30

This was taken from the EORTC QLQ-C30 Scoring Manual [12].

1. Calculate raw score

If items $I_1, I_2, \ldots I_n$ are included in a scale, then the raw score of a scale is calculated as follows:

Raw score = RS = $(I_1 + I_2 + \cdots + I_n)/n$

2. Linear transformation

Apply the linear transformation to 0–100 to obtain the score S,

Functional scales: score = $\{1-(RS-1)/\text{range}\} \times 100$
Symptom scales/items: score = $\{(RS-1)/\text{range}\} \times 100$
Global health status/QOL: score = $\{(RS-1)/\text{range}\} \times 100$

Range is the difference between the maximum possible value of RS and the minimum possible value. The QLQ-C30 has been designed so that all the items in any scale take the same range of values. Therefore, the range of RS equals the range of the item values. Most items are scored 1–4, giving a range = 3. The exceptions are the items contributing to the global health status/QOL, which are 7-point questions with a range = 6, and the initial yes/no item on the earlier versions of the QLQ-C30, which have range = 1.

Appendix 2

See Table 6.

Table 6 Covariance matrix of model 1

	Global health status	Physical functioning	Emotional functioning	Social functioning	Pain symptom	Dyspnoea	Appetite loss	Diarrhea
Global health status	494.11	223.63	253.79	242.87	−276.68	−195.83	−309.79	−13.50
Physical functioning	223.63	414.21	219.61	268.62	−286.63	−275.40	−238.72	−27.33
Emotional functioning	253.79	219.61	549.84	255.38	−252.22	−190.45	−280.17	−48.56
Social functioning	242.87	268.62	255.38	801.87	−293.56	−199.63	−230.29	−92.96
Pain symptom	−276.68	−286.63	−252.22	−293.56	672.78	285.61	319.42	122.34
Dyspnoea	−195.83	−275.40	−190.45	−199.63	285.61	899.59	223.81	28.67
Appetite loss	−309.79	−238.72	−280.17	−230.29	319.42	223.81	1,047.22	155.04
Diarrhea	−13.50	−27.33	−48.56	−92.96	122.34	28.67	155.04	489.75

References

1. Kontodimopoulos, N., Aletras, V. H., Paliouras, D., & Niakas, D. (2009). Mapping the cancer-specific EORTC QLQ-C30 to the preference-based EQ-5D, SF-6D, and 15 D instruments. *Value in Health, 12*(8), 1151–1157.

2. Patrick, D. L., & Deyo, R. A. (1989). Generic and disease-specific measures in assessing health status and quality of life. *Medical Care, 27*(Supp l), S217–S232.

3. Rivero-Arias, O., Quellet, M., Gray, A., Wolstenhome, J., Rothwell, P. M., & Luengo-Fernandez, R. (2009). Mapping the modified rankin scale (mRS) measurement into the generic EuroQol (EQ-5D) health outcome. *Medical Decision Making, 30*(3), 341–354.

4. Cheung, Y. B., Tan, L. C. S., Lau, P. N., Au, W. L., & Luo, N. (2008). Mapping the eight-item Parkinson's Disease Questionnaire (PDQ-8) to the EQ-5D utility index. *Quality of Life Research, 17*, 1173–1181.

5. Bottomley, A., & Aaronson, N. K. (2007). International perspective on health-related quality-of-life research in cancer clinical trials: The European Organization for Research and Treatment of Cancer experience. *Journal of Clinical Oncology, 25*(32), 5082–5086.

6. Aaronson, N. K., Ahmedzai, S., Bergman, B., et al. (1993). The European organization for research and treatment of cancer QLQ C-30: A quality-of-life instrument for use in international clinicaltrials in oncology. *Journal of the National Cancer Institute, 85*(6), 365–376.

7. McKenzie, L., & Pol, M. (2009). Mapping the EORTC QLQ C-30 onto the EQ-5D instrument: The potential to estimate QALYs without generic preference data. *Value in Health, 12*(1), 167–171.

8. Wu, E. Q., Mulani, P., Farrell, M. H., & Sleep, D. (2007). Mapping FACT-P and EORTC QLQ-C30 to patient health status measured by EQ-5D in metastatic hormone-refractory prostate cancer patients. *Value in Health, 10*(5), 408–414.

9. Crott, R., & Briggs, A. (2010). Mapping the QLQ-C30 quality of life cancer questionnaire to EQ-5D patient preferences. *The European Journal of Health Economics, 11*(4), 427–434.

10. Korea National Cancer Information Center. (2010). Prevalence of cancer. cited August 28, 2010, from URL http://www.cancer.go.kr/cms/statistics/stat/1373001_1611.html.

11. Health Insurance Review and Assessment Services. (2006). Korean pharmaceutical economic evaluation guidelines. Cited December 25, 2010, from URL http://www.ispor.org/PEguidelines/source/Korean_PE_Guidelines_Korean_Version.pdf.

12. Fayers, P. M., Aaronson, N. K., Bjordal, K., Curan, D., & Groenvold, M. (1999). *On behalf of EORTC quality of life study group. EORTC QLQ-C30 scoring manual* (2nd ed.). Brussels: EORTC.

13. Sprangers, M. A. G., Groenvold, M., Arraras, J. I., et al. (1996). the EORTC breast cancer-specific quality-of-life questionnaire module: First results from a three-country field study. *Journal of Clinical Oncology, 14*(10), 2756–2768.

14. The EuroQol Group. EQ-5D. Cited December 25, 2010, from URL http://www.euroqol.org.

15. Lee, Y. K., Nam, H. S., Chuang, L. H., et al. (2009). South Korean time trade-off values for EQ-5D health states: Modeling with observed values for 101 health states. *Value in Health, 12*(8), 1187–1193.

16. Borras, J. M., Sanchez-Hernandez, A., Navarro, M., et al. (2001). Compliance, satisfaction, and quality of life of patients with colorectal cancer receiving home chemotherapy or outpatient treatment: A randomized controlled trial. *British Medical Journal, 322*, 1–5.

17. World Health Organization. (1980). WHO handbook for reporting results of cancer treatment. *Neoplasma, 20*, 37–46.

18. Dakin, H., Petrou, S., Haggard, M., Benge, S., & Williamson, I. (2010). Mapping analyses to estimate health utilities based on responses to the OM8–30 otitis media questionnaire. *Quality of Life Research, 19*(1), 65–80.

19. Versteegh, M. M., Rowen, D., Brazier, J. B., & Stolk, E. A. (2010). Mapping onto EQ-5 D for patients in poor health. *Health and Quality of Life Outcomes, 8*, 141.

20. Brazier, J. E., Yang, Y., Tsuchiya, A., & Rowen, D. L. (2009). A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *European Journal of Health Economics, 2010*(11), 215–225.

21. Rowen, D., Brazier, J., & Roberts, J. (2009). Mapping SF-36 onto the EQ-5 D index: How reliable is the relationship? *Health and Quality of Life Outcomes, 7*, 27.

22. Sauerland, S., Weiner, S., Dolezalova, K., Angrisani, L., Noguera, C. M., Garcia-Caballero, M., et al. (2009). Mapping utility scores from a disease-specific quality-of-life measure in Bariatric surgery patients. *Value in Health, 12*(2), 364–370.

23. Vittinghoff, E., Shiboski, S. C., Glidden, D. V., McCulloch, C. E. (2005). *Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models* (Chap. 5, pp. 147–148). New York: Springer Science+Business Media, Inc.

24. SAS/STAT(R) 9.2 (2009). User's Guide (2nd ed.), (Chap. 4, pp. 98). Cary, North Carolina, USA: SAS Institute Inc.

25. Korea Center for Disease Control and Prevention. (2007). The Korea national health and nutrition examination survey (KNHANES IV). Seoul: Korea Ministry of Health, Welfare, and Family Affairs, p. 290.