



# Mapping the Chinese Version of the EORTC QLQ-BR53 Onto the EQ-5D-5L and SF-6D Utility Scores

Tong Liu<sup>1,2</sup> · Shunping Li<sup>1,2</sup> · Min Wang<sup>3</sup> · Qiang Sun<sup>1,2</sup> · Gang Chen<sup>4</sup>

Published online: 8 May 2020  
© Springer Nature Switzerland AG 2020

## Abstract

**Objective** This study aimed to develop mapping algorithms from the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-BR53, including EORTC QLQ-C30 and QLQ-BR23) onto the 5-level EQ-5D (EQ-5D-5L) and Short Form 6D (SF-6D) utility scores.

**Methods** The data were taken from 607 breast cancer patients in mainland China. The EQ-5D-5L and SF-6D instruments were scored using Chinese-specific tariffs. Three model specifications and seven statistical techniques were used to derive mapping algorithms, including ordinary least squares (OLS), Tobit, censored least absolute deviation (CLAD) model, generalized linear model (GLM), robust MM-estimator, finite mixtures of beta regression model for directly estimating health utility, and using ordered logit regression (OLOGIT) to predict response levels. A five-fold cross-validation approach was conducted to test the generalizability of each model. Two key goodness-of-fit statistics (mean absolute error and mean squared error) and three secondary statistics were employed to choose the optimal models.

**Results** Participants had a mean  $\pm$  standard deviation (SD) age of  $49.0 \pm 9.8$  years. The mean  $\pm$  SD health state utility scores were  $0.828 \pm 0.184$  (EQ-5D-5L) and  $0.646 \pm 0.125$  (SF-6D). Mapping performance was better when both the QLQ-C30 and QLQ-BR23 dimensions were considered rather than when either of these dimensions were used alone. The mapping functions from the optimal direct mapping and indirect mapping approaches were reported.

**Conclusions** The algorithms reported in this paper enable EORTC QLQ-BR53 breast cancer data to be mapped into utilities predicted from the EQ-5D-5L and SF-6D. The algorithms allow for the calculation of quality-adjusted life years for use in breast cancer cost-effectiveness analyses studies.

## Key Points for Decision Makers

The goodness-of-fit performance of mapping onto the Short Form 6D (SF-6D) is significantly better than using the 5-level EQ-5D (EQ-5D-5L), thus the derived SF-6D utilities from mapping functions are more appropriately for calculating quality-adjusted life year.

The response mapping used in this study can largely improve prediction accuracy, especially for SF-6D.

It should be noted that the best models were not able to predict negative EQ-5D-5L utility scores, hence they may not predict well for the poor health states.

**Electronic Supplementary Material** The online version of this article (<https://doi.org/10.1007/s40271-020-00422-x>) contains supplementary material, which is available to authorized users.

✉ Shunping Li  
lishunping@sdu.edu.cn

<sup>1</sup> School of Health Care Management, Cheeloo College of Medicine, Shandong University, Jinan 250012, China

<sup>2</sup> NHC Key Laboratory of Health Economics and Policy Research (Shandong University), Jinan 250012, China

<sup>3</sup> Qingdao Municipal Hospital, Qingdao 266011, China

<sup>4</sup> Flinders Centre for Innovation in Cancer, Flinders University, Adelaide, SA 5042, Australia

## 1 Introduction

Breast cancer is the most common and second deadliest cancer in women worldwide, and it is estimated that 1.4 million women a year receive a diagnosis of breast cancer [1, 2]. In 2018, it was the most prevalent cancer and the fifth leading cause of cancer death in Chinese women [3]. The treatment for breast cancer patients resulted in a substantial financial burden on the Chinese health care system [4].

To prioritize health care resource allocation, health economic evaluation, especially cost-utility analysis (CUA) has become a preferred method [5]. In CUA, effectiveness is measured by using quality-adjusted life years (QALYs), which are calculated by multiplying the life-years by the health state utility scores. The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ)-BR53 (which consists of QLQ-C30 and QLQ-BR23) is one of the most widely used disease-specific outcome measures in breast cancer studies [6–8]; however, it is not a preference-based instrument and cannot be used to calculate the QALYs.

Mapping (or ‘crosswalk’) provides a solution to predict health state utility scores from a non-preference-based quality of life instrument [9]. The predicted utility values can then be analyzed using standard methods for trial-based analyses, or summarized for each health state within an economic model [10, 11]. This method has successfully been used in predicting the 3-level EQ-5D utility using breast cancer patients [12, 13]. This study aimed to develop mapping algorithms from the QLQ-BR53 onto either the 5-level EQ-5D (EQ-5D-5L) or the Short Form 6D (SF-6D) utility scores based on breast cancer patients in China. The output from this study will facilitate future CUAs in which either the QLQ-C30, QLQ-BR23, or 53-item QLQ-BR53 is included.

## 2 Materials and Methods

### 2.1 Study Population

A total of 621 female inpatients with breast cancer were recruited from Qingdao Municipal Hospital in China between October 2014 and February 2015. The inclusion criteria were patients who had been diagnosed with breast cancer, aged 18 years and older, and provided written consent to participate in the study before the interview. The exclusion criteria were patients who were unwilling to provide informed consent or could not understand the questionnaires, had breast cancer in combination with other serious diseases, or the patient was not yet 18 years of age at the time of the survey. Information on sociodemographic

characteristics, clinical data, and health-related quality of life (HRQoL) was collected using two methods in a single visit: (1) face-to-face interview using an essential information questionnaire as well as three standard instruments, i.e. QLQ-BR53, EQ-5D-5L and SF-36 (used to derive the SF-6D); and (2) using the medical records of patients for clinical information. Among 621 respondents, 14 patients who had missing values on key questions were excluded from the mapping analysis. For the remaining 607 patients, there were some missing values in four QLQ-BR53 domain scores. Informed consent was obtained from all participants before the completion of the instrument survey. Ethical approval (reference no. 20131002) was obtained from the Ethics Review Board of the School of Public Health, Shandong University, and the research adhered to the tenets of the Declaration of Helsinki.

### 2.2 Instruments

#### 2.2.1 QLQ-BR53

The QLQ-BR53 consists of two instruments: (1) the QLQ-C30, which contains 30 items and covers five functional scales, three symptom scales, a global health status/QOL scale, and six single items; and (2) the QLQ-BR23, which contains 23 items divided into five multi-item scales assessing systemic therapy side effects, arm symptoms, breast symptoms, body image and sexual functioning, besides, three single items evaluating sexual enjoyment, hair loss and future perspectives [6]. The score for each subscale was calculated by summing responses for all items in each subscale according to the official EORTC scoring manual [14]. The raw scores of the participants’ responses were then linearly transformed to a 0–100 scale, with a higher score indicating a better quality of life for the functioning and global health status, but a poorer quality of life for severe symptomatic problems. The QLQ-BR53 shows reasonable reliability, validity and responsiveness, and can be used to measure quality of life for Chinese patients with breast cancer [7].

#### 2.2.2 EQ-5D-5L

The EQ-5D-5L is a common and validated instrument consisting of five dimensions of health (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), which are characterized by five levels (no problems, slight problems, moderate problems, severe problems, and extreme problems) [15]. The EQ-5D-5L was scored using a Chinese-specific tariff developed based on a time trade-off method. The Chinese tariff has a theoretical range of scores from –0.149 to 1.0 [16].

### 2.2.3 SF-6D

The SF-6D was constructed from 11 items selected from the SF-36 [17]. The SF-6D is based on a six-dimensional health state classification that assesses physical functioning, role limitations, social functioning, bodily pain, mental health, and vitality. Each dimension of the SF-6D has 4–6 levels and can be used to describe 18,000 health states [18]. In the absence of the mainland China utility algorithm, the Hong Kong tariff was used for this study, and has a theoretical range of scores from 0.315 to 1.0 [19].

## 2.3 Statistical Analysis

### 2.3.1 Crosswalks

This study was conducted in accordance with the ‘Mapping onto Preference-based measures reporting Standards’ (MAPS) checklist [20] (see Online Resource 1) and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Good Practices for Outcome Research Task Force report on mapping to estimate health state utilities [9], as well as a systematic review on mapping studies in the annual report of the National Institute for Health and Care Excellence (NICE) [21].

Patient characteristics were described using mean  $\pm$  standard deviation (SD) or percentage in the sample. We tested for normality of variables using the Shapiro–Wilks test. The degree of conceptual overlap between the source and the target variables was examined using Spearman’s rank correlation. Three model specifications were considered in this study. Considering the potential multicollinearity among the QLQ-BR53 items in the regression, the mapping functions focused on the dimension scores of the QLQ-BR53. In Model 1, only dimensions from QLQ-C30 were considered, while in Model 2 only dimensions from QLQ-BR23 were considered, and, finally, in Model 3, all dimensions from QLQ-C30 and QLQ-BR23 were considered. One additional and commonly available demographic characteristic from other datasets, i.e. age, was also considered in the mapping functions. The potential non-linear effects from dimension scores can be captured by some econometric methods adopted and introduced below. A stepwise regression technique was used to help choose the final statistically significant (i.e.  $p < 0.05$ ) predictors.

Both direct and indirect (response) mapping analyses were conducted. In this study, six commonly used statistical methods were adopted for direct mapping and one method was used for indirect mapping [22]. It should be noted that when mapping onto the SF-6D utility, two censored models were not used since there was no censoring issue for the SF-6D utility.

**Table 1** Patient characteristics ( $N=607$ )

Characteristics	
Age (years) [mean (SD)]	49.0 $\pm$ 9.8
Registered residence	
Urban area	308 (50.7)
Rural area	299 (49.3)
Marital status	
Married	538 (88.6)
Single	27 (4.4)
Divorced/widowed	42 (7.0)
Education level	
Primary school /illiteracy	199 (32.8)
Junior middle school	128 (21.1)
Senior high school/technical secondary school	131 (21.6)
College degree and higher	149 (24.5)
Type of health insurance	
Out-of-pocket payments	24 (4)
Medical insurance	583 (96)
Household income (10,000 yuan)	
< 3	160 (26.4)
3–8	262 (43.2)
9–15	153 (25.2)
> 15	32 (5.2)
Disease course (months)	
< 13	189 (31.1)
13–36	197 (32.5)
37–60	112 (18.5)
> 60	109 (17.9)
TNM stages	
0–I	174 (28.6)
II	142 (23.4)
III	218 (35.9)
IV	73 (12.1)
Menopausal status	
Premenopausal	334 (55.0)
Postmenopausal	273 (45.0)
Disease status classification	
Primary breast cancer within 1 year	156 (25.7)
Recurrence of breast cancer within 1 year	65 (10.7)
Primary breast cancer or recurrence for more than 1 year	245 (40.4)
Metastasis for more than 1 year	141 (23.2)

Data are expressed as  $n$  (%) unless otherwise specified  
SD standard deviation

1. The ordinary least square (OLS) model was used to estimate the unknown parameters by minimizing the sum of squared errors from the data [23]. This is the most widely used method in mapping studies [24].
2. The Tobit model takes a better account of the censored nature of EQ-5D data, dealing with truncated data, and

**Table 2** Descriptive statistics for health utility and HRQoL variables

Variable	N	Mean (SD)	Median (IQR)
EQ-5D-5L	607	0.828 (0.184)	0.882 (0.734–1)
SF-6D	607	0.646 (0.125)	0.62 (0.537–0.774)
EORTC QLQ-C30			
Physical functioning (PF score)	607	75.464 (17.252)	80.000 (66.667–86.667)
Role functioning (RF score)	607	77.375 (25.561)	83.333 (66.667–100.000)
Emotional functioning (EF score)	607	74.218 (19.720)	75.000 (66.667–91.667)
Cognitive functioning (CF score)	607	76.853 (19.485)	83.333 (66.667–83.333)
Social functioning (SF score)	607	69.852 (24.596)	66.667 (66.667–100.000)
Fatigue (FA score)	607	34.047 (18.057)	33.333 (22.222–44.444)
Nausea and vomiting (NV score)	607	19.028 (21.472)	16.667 (0.000–33.333)
Pain (PA score)	607	28.913 (19.835)	33.333 (16.667–33.333)
Dyspnoea (DY score)	607	17.188 (22.233)	0.000 (0.000–33.333)
Insomnia (SL score)	607	31.466 (24.377)	33.333 (0.000–33.333)
Appetite loss (AP score)	607	24.108 (25.314)	33.333 (0.000–33.333)
Constipation (CO score)	607	24.657 (26.357)	33.333 (0.000–33.333)
Diarrhoea (DI score)	607	10.434 (18.886)	0.000 (0.000–33.333)
Financial difficulties (FI score)	607	34.706 (28.659)	33.333 (0.000–33.333)
Global health status (QL score)	607	53.817 (14.673)	50.000 (50.000–66.667)
QLQ-BR23			
Body image (BI score)	607	64.896 (25.008)	66.667 (50.000–75.000)
Sexual functioning (SEF score)	604	88.963 (15.933)	100.000 (66.667–100.000)
Sexual enjoyment (SEE score)	598	88.294 (19.593)	100.000 (66.667–100.000)
Future perspective (FU score)	607	51.455 (31.425)	66.667 (33.333–66.667)
Systemic therapy side effects (ST score)	606	24.745 (16.870)	23.810 (9.524–33.333)
Breast symptoms (BS score)	607	17.147 (19.817)	8.333 (0.000–33.333)
Arm symptoms (AS score)	607	20.245 (19.658)	22.222 (0.000–33.333)
Upset by hair loss (HL score)	592	38.626 (30.261)	33.333 (33.333–66.667)

HRQoL health-related quality of life, SD standard deviation, IQR interquartile range

can approximate for skewed data by setting the upper limit to 1 [25].

3. The censored least absolute deviation (CLAD) model is a censored model to estimate conditional medians, such that it is robust to distributional assumption and heteroscedasticity [26].
4. The generalized linear model (GLM) allows for the non-normal distribution of dependent variables (e.g. left/negatively skewed utility scores) [27]. In this study, the GLM was estimated using a Gaussian distribution with a log-link function, which was identified to produce better goodness-of-fit from different combinations of family and link functions.
5. The robust MM-estimator (MM) model is designed to deal with some limitations of traditional regression methods, including heteroscedasticity and the presence of outliers. It has been shown to have both a high breakdown point (i.e. the percentage of incorrect observations an estimator can handle before giving an incorrect result) and a high efficiency [28]. It was first introduced into the mapping literature by Chen and colleagues for both adolescent and adult samples, and was found to have good performance [29, 30].
6. The finite mixtures of beta regression model (BETA-MIX) is a general version of the truncated inflated beta regression model for variables with truncated supports either at the top or bottom of the distribution [31]. The model is robust to skewness and can estimate both unimodal and bimodal utilities [32].

The choice of the above methods cover a wide range of potential challenges when deriving the mapping functions. For instance, the MM and CLAD are robust estimators that are less influenced by potential outliers in the dataset and they can cope with the potential heteroscedasticity [29]. The Tobit and CLAD models can cope with the censoring issue that a large proportion of health state utilities equal 1 in the dataset [33]. The GLM and BETAMIX estimators are usually handling the skewness distribution of the study data [34]. Although, theoretically, different estimators have their strengths and may be better suited for the study data, the empirical evidence is crucial for the justification

**Table 3** Spearman rank correlations of QLQ-BR53 dimension scores with EQ-5D-5L and SF-6D utility values ( $N=607$ )

QLQ-BR53	EQ-5D-5L						SF-6D						
	MO	SC	UA	PD	AD	Utility	PH	RL	SF	PA	MH	VT	Utility
Physical functioning (PF score)	<b>0.62</b>	<b>0.62</b>	<b>0.61</b>	0.47	0.40	<b>0.64</b>	<b>0.62</b>	0.49	0.04	<b>0.50</b>	0.26	0.28	<b>0.62</b>
Role functioning (RF score)	<b>0.56</b>	<b>0.60</b>	<b>0.58</b>	0.49	0.39	<b>0.61</b>	<b>0.58</b>	<b>0.51</b>	0.15	<b>0.52</b>	0.22	0.32	<b>0.63</b>
Emotional functioning (EF score)	0.45	0.47	0.49	0.42	0.46	<b>0.55</b>	0.49	0.43	0.05	0.41	0.22	0.26	<b>0.55</b>
Cognitive functioning (CF score)	0.47	0.49	<b>0.54</b>	0.43	0.37	<b>0.56</b>	<b>0.51</b>	0.44	0.07	0.44	0.22	0.29	<b>0.55</b>
Social functioning (SF score)	0.41	0.45	0.46	0.41	0.36	0.49	0.44	0.39	0.22	0.43	0.19	0.34	<b>0.52</b>
Fatigue (FA score)	0.43	0.46	0.48	0.47	0.42	<b>0.55</b>	0.44	0.41	0.02	0.45	0.29	0.28	0.48
Nausea and vomiting (NV score)	0.48	0.48	0.49	0.38	0.31	<b>0.50</b>	0.45	0.39	0.01	0.37	0.11	0.26	0.49
Pain (PA score)	0.46	0.48	<b>0.51</b>	<b>0.51</b>	0.45	<b>0.57</b>	0.45	0.43	0.09	<b>0.50</b>	0.16	0.33	<b>0.52</b>
Dyspnoea (DY score)	0.47	0.45	<b>0.50</b>	0.36	0.34	<b>0.51</b>	0.41	0.38	0.03	0.35	0.20	0.18	0.43
Insomnia (SL score)	0.25	0.25	0.27	0.26	0.21	0.31	0.29	0.21	0.10	0.20	0.11	0.21	0.31
Appetite loss (AP score)	0.42	0.41	0.42	0.33	0.28	0.45	0.44	0.36	0.06	0.34	0.11	0.23	0.45
Constipation (CO score)	0.24	0.23	0.28	0.18	0.11	0.25	0.26	0.11	0.05	0.11	0.02	0.12	0.20
Diarrhoea (DI score)	0.38	0.37	0.35	0.21	0.17	0.34	0.38	0.25	0.09	0.19	0.11	0.18	0.36
Financial difficulties (FI score)	0.33	0.39	0.39	0.32	0.26	0.40	0.39	0.32	0.22	0.34	0.20	0.28	0.45
Global health status (QL score)	0.09	0.14	0.14	0.17	0.14	0.16	0.08	0.02	0.26	0.15	0.07	0.22	0.10
Body image (BI score)	0.37	0.35	0.37	0.37	0.39	0.44	0.29	0.38	0.06	0.33	0.19	0.22	0.39
Sexual functioning (SEF score)	0.28	0.25	0.25	0.15	0.11	0.25	0.36	0.25	0.01	0.17	0.03	0.12	0.34
Sexual enjoyment (SEE score)	0.28	0.25	0.27	0.13	0.14	0.27	0.33	0.21	0.07	0.13	0.02	0.10	0.29
Future perspective (FU score)	0.23	0.25	0.26	0.22	0.31	0.30	0.25	0.27	0.13	0.24	0.15	0.18	0.30
Systemic therapy side effects (ST score)	0.48	0.49	<b>0.51</b>	0.46	0.40	<b>0.56</b>	0.44	0.42	0.11	0.41	0.15	0.31	<b>0.51</b>
Breast symptoms (BS score)	0.47	<b>0.52</b>	<b>0.50</b>	0.42	0.41	<b>0.55</b>	<b>0.51</b>	0.48	0.02	0.42	0.24	0.27	<b>0.56</b>
Arm symptoms (AS score)	0.46	<b>0.53</b>	<b>0.54</b>	<b>0.50</b>	0.48	<b>0.58</b>	0.49	<b>0.54</b>	0.11	0.48	0.20	0.27	<b>0.60</b>
Upset by hair loss (HL score)	0.26	0.25	0.27	0.35	0.37	0.37	0.21	0.33	0.09	0.32	0.16	0.21	0.33

All correlations are statistically significant at  $<1\%$ . For all measures, the highest value is the best possible outcome, and strong correlations ( $\geq 0.5$ ) are highlighted in bold

MO mobility, SC self-care, UA usual activities, PD pain/discomfort, AD anxiety/depression, PH physical health, RL role limitation, SF social functioning, PA pain, MH mental health, VT vitality

of the optimal estimator in the mapping studies. Although its assumptions have been violated, the OLS still has better performance when compared with other advanced methods [10, 11].

For indirect mapping, the response levels of each dimension were firstly predicted, followed by using the country-specific tariff to generate the overall utility score. Since the response level is an ordinal variable (e.g. no problems, slight problems, moderate problems, severe problems, and extreme problems), a multinomial or ordered logit model (OLOGIT) is commonly used [35]. Following on from Chen et al. [36], an OLOGIT [37] was adopted in this study and was followed by using the corresponding Chinese-specific algorithms to calculate the overall EQ-5D-5L/SF-6D utility scores.

### 2.3.2 Goodness-of-Fit Indicators

Predictive ability was mainly assessed based on the mean absolute error (MAE) and the mean squared error (MSE). Three additional indicators were also considered, including

Lin's concordance correlation coefficient (CCC), and the proportion of predicted utilities deviating from observed values by absolute error  $>0.05$  and  $>0.1$ . With the exception of CCC (the higher the value, the better the performance), the lower the value, the better the mapping performance, as indicated by goodness-of-fit indicators. In cases where the prediction exceeds the theoretical range of the targeting utility (e.g. the predicted maximum utility is above 1.0), to mimic the real-life solution, those predictions were truncated at the theoretical maximum and/or minimum utility scores before the goodness-of-fit statistics were calculated. The goodness-of-fit results without this adjustment can be found in Online Resource 1.

Two internal validation procedures were performed. In the first (Validation I), the whole sample was randomly divided into five groups. In each group, 80% of the sample was used to calculate the mapping algorithm and the remaining 20% was used to predict the health state utility with the above mapping algorithm. These procedures were repeated five times, such that all groups were used as both predictors

Table 4 Goodness-of-fit results from validation analyses

	Validation I					Validation II						
	Pooled sample (N=607)					Random sample (N=300)						
	(1) Mean	(2) MAE	(3) MSE	(4) CCC	(5) AE>0.05 (%)	(6) AE>0.1 (%)	(7) Mean	(8) MAE	(9) MSE	(10) CCC	(11) AE>0.05 (%)	(12) AE>0.1 (%)
Observed (EQ-5D-5L)	0.8278	-	-	-	-	-	0.8230	-	-	-	-	-
Observed (SF-6D)	0.6459	-	-	-	-	-	0.6397	-	-	-	-	-
<i>Model 1</i>												
<b>EQ-5D-5L</b>												
OLS <sup>a</sup>	0.8301	0.1027	0.2053	0.605	67.87	37.66	0.8223	0.0981	0.1962	0.700	63.00	37.33
Tobit <sup>a</sup>	0.8550	<b>0.0983</b>	<b>0.1965</b>	<b>0.649</b>	<b>59.31</b>	<b>36.35</b>	0.8487	0.0970	0.1940	0.718	<b>60.33</b>	36.33
CLAD <sup>a</sup>	0.8525	0.1097	0.2194	0.554	67.38	44.15	0.8427	0.0945	0.1889	0.692	61.67	37.00
GLM <sup>a</sup>	0.8319	0.1024	0.2048	0.585	65.90	38.16	0.8232	0.1011	0.2022	0.641	64.00	36.67
MM <sup>a</sup>	0.8546	0.1017	0.2034	0.570	67.05	37.17	0.8539	0.0984	0.1967	0.581	63.33	<b>33.33</b>
BETAMIX	0.8502	0.1068	0.2135	0.612	65.40	37.34	0.8309	0.1026	0.2052	0.735	65.67	34.33
OLOGIT	0.8672	0.1006	0.2013	0.626	64.25	42.17	0.8610	<b>0.0909</b>	<b>0.1819</b>	<b>0.748</b>	61.33	38.00
<b>SF-6D</b>												
OLS <sup>b</sup>	0.6449	0.0773	0.1546	0.638	69.41	28.95	0.6400	0.0724	0.1447	0.708	64.33	23.00
GLM	0.6445	0.0743	0.1486	0.660	66.28	25.16	0.6395	0.0693	0.1385	0.729	59.67	<b>21.67</b>
MM <sup>b</sup>	0.6432	0.0789	0.1578	0.646	67.43	32.73	0.6445	0.0710	0.1420	0.716	60.67	23.67
BETAMIX	0.6440	0.0768	0.1535	0.637	66.78	28.95	0.6422	0.0719	0.1439	0.728	62.67	25.00
OLOGIT	0.6617	<b>0.0656</b>	<b>0.1312</b>	<b>0.712</b>	<b>43.66</b>	<b>25.04</b>	0.6568	<b>0.0651</b>	<b>0.1302</b>	<b>0.730</b>	<b>44.00</b>	26.33
<i>Model 2</i>												
<b>EQ-5D-5L</b>												
OLS <sup>a</sup>	0.8313	0.0983	0.1965	0.571	60.13	36.84	0.8224	0.1042	0.2084	0.588	64.00	41.33
Tobit <sup>a</sup>	0.8610	0.0973	0.1946	<b>0.597</b>	55.85	36.51	0.8484	0.1019	0.2037	0.626	62.00	38.67
CLAD <sup>a</sup>	0.8525	<b>0.0949</b>	<b>0.1899</b>	0.548	<b>54.20</b>	31.30	0.8606	<b>0.1007</b>	<b>0.2014</b>	0.506	<b>54.33</b>	34.00
GLM <sup>a</sup>	0.8309	0.0991	0.1981	0.549	59.80	36.68	0.8231	0.1048	0.2095	0.543	62.00	40.33
MM <sup>a</sup>	0.8656	0.0970	0.1939	0.510	57.99	<b>30.76</b>	0.8570	0.1026	0.2051	0.457	57.33	<b>31.67</b>
BETAMIX	0.8356	0.1035	0.2070	0.574	62.11	31.25	0.8186	0.1069	0.2138	<b>0.636</b>	68.67	34.33
OLOGIT	0.8437	0.1155	0.2310	0.435	59.31	38.71	0.7791	0.1383	0.2765	0.500	66.67	50.33
<b>SF-6D</b>												
OLS	0.6468	0.0816	0.1632	0.579	69.74	31.09	0.6416	0.0803	0.1606	0.585	62.67	35.00
GLM	0.6457	0.0803	0.1606	0.598	65.30	<b>29.65</b>	0.6413	0.0795	0.1591	0.581	60.67	33.00
MM <sup>b</sup>	0.6459	<b>0.0786</b>	<b>0.1572</b>	<b>0.605</b>	64.47	32.07	0.6439	<b>0.0780</b>	<b>0.1560</b>	<b>0.605</b>	57.33	32.33
BETAMIX	0.6488	0.0804	0.1609	0.581	65.63	31.25	0.6433	0.0827	0.1654	0.575	64.67	31.67
OLOGIT	0.64413	0.0787	0.1573	0.586	<b>50.74</b>	29.77	0.6358	0.0843	0.1685	0.550	<b>50.67</b>	<b>31.00</b>

Table 4 (continued)

	Validation I					Validation II						
	Pooled sample (N=607)					Random sample (N=300)						
	(1) Mean MAE	(2) MAE	(3) MSE	(4) CCC	(5) AE>0.05 (%)	(6) AE>0.1 (%)	(7) Mean	(8) MAE	(9) MSE	(10) CCC	(11) AE>0.05 (%)	(12) AE>0.1 (%)
<i>Model 3</i>												
<b>EQ-5D-5L</b>												
OLS <sup>a</sup>	0.8280	0.0992	0.1985	0.617	61.61	37.01	0.8224	0.0937	0.1874	0.722	60.33	35.67
Tobit <sup>a</sup>	0.8586	0.0981	0.1963	0.591	59.31	36.84	0.8462	0.0925	0.1850	<b>0.740</b>	58.67	37.00
CLAD <sup>a</sup>	0.8561	<b>0.0959</b>	<b>0.1917</b>	<b>0.653</b>	61.45	33.61	0.8310	<b>0.0888</b>	<b>0.1775</b>	0.716	<b>53.67</b>	35.00
GLM <sup>a</sup>	0.8276	0.0983	0.1966	0.608	64.25	36.68	0.8243	0.0966	0.1931	0.661	62.00	37.00
MM <sup>a</sup>	0.8678	0.0966	0.1931	0.528	<b>56.67</b>	<b>29.93</b>	0.8544	0.0931	0.1862	0.579	54.67	<b>30.67</b>
BETAMIX	0.8491	0.1041	0.2083	0.604	58.32	34.38	0.8281	0.1004	0.2008	0.721	56.00	34.67
OLOGIT	0.8481	0.1056	0.2112	0.559	61.78	39.54	0.8409	0.0973	0.1946	0.718	59.67	38.33
<b>SF-6D</b>												
OLS <sup>b</sup>	0.6424	0.0729	0.1458	0.676	63.98	<b>23.03</b>	0.6431	0.0686	0.1371	0.721	59.67	20.00
GLM	0.6432	0.0701	0.1401	0.696	58.88	23.52	0.6426	0.0662	0.1325	0.739	56.00	<b>19.33</b>
MM <sup>b</sup>	0.6482	0.0714	0.1428	0.693	58.22	26.48	0.6470	0.0668	0.1337	0.735	54.67	21.33
BETAMIX	0.6456	0.0739	0.1479	0.668	62.83	26.81	0.6446	0.0708	0.1416	0.725	57.33	24.00
OLOGIT	0.6552	<b>0.0648</b>	<b>0.1296</b>	<b>0.711</b>	<b>43.16</b>	24.38	0.6445	<b>0.0610</b>	<b>0.1224</b>	<b>0.747</b>	<b>39.33</b>	23.33

OLS ordinary least square, CLAD censored least absolute deviation, GLM generalized linear model, MM robust MM-estimator, BETAMIX the finite mixtures of beta regression model, OLOGIT ordered logit regression, MAE mean absolute error, MSE mean squared error, CCC Lin's concordance correlation coefficient, AE > 0.05 and AE > 0.1 signify the proportion of predicted utilities deviating from observed values by absolute error > 0.05 and > 0.1, respectively

The best results among the three direct mapping methods are highlighted in bold

<sup>a</sup>The adjusted goodness-of-fit results by specifying the maximum predicted utility score to be 1

<sup>b</sup>The adjusted goodness-of-fit results by specifying the minimum predicted utility score to be 0.315



**Table 5** Goodness-of-fit results from the full sample ( $N=607$ )

Models	(1) Mean	(2) Minimum	(3) Maximum	(4) MAE	(5) MSE	(6) CCC	(7) AE > 0.05 (%)	(8) AE > 0.1 (%)
Observed (EQ-5D-5L)	0.8278	-0.3	1	-	-	-	-	-
Observed (SF-6D)	0.6459	0.315	0.962	-	-	-	-	-
<i>Model 1</i>								
<b>EQ-5D-5L</b>								
OLS <sup>a</sup>	0.8275	0.2140	1	0.0915	0.1830	0.687	59.14	34.10
Tobit <sup>a</sup>	0.8544	0.1099	1	0.0994	0.1988	<b>0.707</b>	62.93	33.28
CLAD <sup>a</sup>	0.8416	0.1927	1	0.0893	0.1787	0.693	<b>58.48</b>	32.29
GLM <sup>a</sup>	0.8279	0.3823	1	0.0924	0.1849	0.657	58.65	34.43
MM <sup>a</sup>	0.8555	0.3643	1	0.0927	0.1854	0.612	61.12	<b>30.64</b>
BETAMIX	0.8380	-0.1576	0.9782	0.0993	0.1987	0.682	61.29	33.11
OLOGIT	0.8592	-0.1620	1	<b>0.0883</b>	<b>0.1766</b>	0.703	58.81	35.75
<b>SF-6D</b>								
OLS <sup>b</sup>	0.6460	0.3150	0.7692	0.0733	0.1467	0.675	65.90	26.03
GLM	0.6458	0.3372	0.7908	0.0702	0.1405	0.693	60.30	<b>23.23</b>
MM <sup>b</sup>	0.6495	0.3150	0.7834	0.0718	0.1436	0.689	62.60	23.89
BETAMIX	0.6476	0.3440	0.8124	0.0720	0.1440	0.690	64.42	25.04
OLOGIT	0.6623	0.3380	0.8340	<b>0.0641</b>	<b>0.1283</b>	<b>0.717</b>	<b>40.53</b>	24.38
<i>Model 2</i>								
<b>EQ-5D-5L</b>								
OLS <sup>a</sup>	0.8273	0.2854	1	0.0942	0.1884	0.613	60.79	32.45
Tobit <sup>a</sup>	0.8558	0.1952	1	0.1053	0.2106	<b>0.643</b>	65.24	34.93
CLAD <sup>a</sup>	<b>0.8544</b>	<b>0.3149</b>	<b>1</b>	<b>0.0887</b>	<b>0.1773</b>	0.600	<b>49.75</b>	30.31
GLM <sup>a</sup>	0.8271	0.4360	1	0.0939	0.1877	0.599	58.81	32.62
MM <sup>a</sup>	0.8620	0.3909	1	0.0908	0.1815	0.542	52.55	<b>27.84</b>
BETAMIX	0.8308	-0.1466	0.9659	0.1002	0.2004	0.626	60.96	31.14
OLOGIT	0.8474	-0.182	1	0.1075	0.2149	0.488	55.02	37.07
<b>SF-6D</b>								
OLS	0.6459	0.3315	0.7618	0.0767	0.1533	0.618	65.90	27.68
GLM	0.6458	0.3785	0.7691	0.0747	0.1494	0.629	62.77	<b>26.85</b>
MM <sup>b</sup>	0.6485	0.3150	0.7807	<b>0.0742</b>	<b>0.1485</b>	<b>0.646</b>	<b>57.17</b>	28.50
BETAMIX	0.6462	0.3628	0.7768	0.0764	0.1529	0.620	63.76	28.34
OLOGIT	0.4937	0.3360	0.6290	0.1555	0.3110	0.191	86.00	59.14
<i>Model 3</i>								
<b>EQ-5D-5L</b>								
OLS <sup>a</sup>	0.8272	0.1688	1	0.0875	0.1750	0.713	57.00	32.13
Tobit <sup>a</sup>	0.8527	0.1057	1	0.0984	0.1967	<b>0.722</b>	62.44	34.27
CLAD <sup>a</sup>	<b>0.8578</b>	<b>0.3065</b>	<b>1</b>	<b>0.0839</b>	<b>0.1679</b>	0.657	<b>51.57</b>	27.51
GLM <sup>a</sup>	0.8272	0.3827	1	0.0885	0.1771	0.683	57.33	31.30
MM <sup>a</sup>	0.8636	0.3525	1	0.0869	0.1739	0.593	51.89	<b>26.19</b>
BETAMIX	0.8369	-0.2082	0.9767	0.0966	0.1932	0.697	54.20	31.96
OLOGIT	0.8400	-0.104	1	0.0990	0.1980	0.587	55.35	35.09
<b>SF-6D</b>								
OLS <sup>b</sup>	0.6461	0.3150	0.7783	0.0675	0.1350	0.716	61.29	21.42
GLM	0.6458	0.3255	0.7998	0.0655	0.1310	0.731	56.51	<b>21.09</b>
MM <sup>b</sup>	0.6511	0.3150	0.7879	0.0666	0.1331	0.725	54.04	22.08
BETAMIX	0.6476	0.3425	0.8264	0.0677	0.1354	0.724	58.65	21.42
OLOGIT	0.6554	0.3360	0.8340	<b>0.0610</b>	<b>0.1219</b>	<b>0.742</b>	<b>38.22</b>	22.41

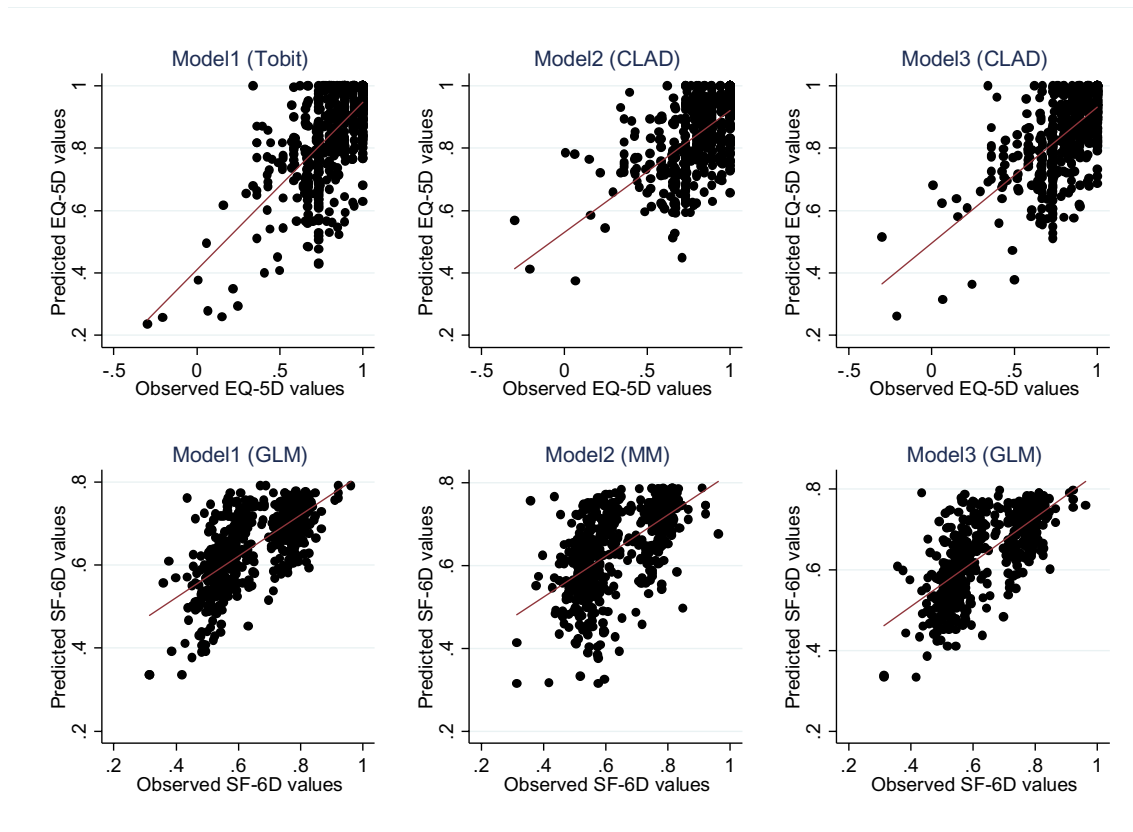
OLS ordinary least square, CLAD censored least absolute deviation, GLM generalized linear model, MM robust MM-estimator, BETAMIX the finite mixtures of beta regression model, OLOGIT ordered logit regression, MAE mean absolute error, MSE mean squared error, CCC Lin's concordance correlation coefficient, AE > 0.05 and AE > 0.1 signify the proportion of predicted utilities deviating from observed values by absolute error > 0.05 and > 0.1, respectively

Best results among the three direct mapping methods are highlighted in bold

<sup>a</sup>The adjusted goodness-of-fit results by specifying the maximum predicted utility score to be 1

<sup>b</sup>The adjusted goodness-of-fit results by specifying the minimum predicted utility score to be 0.315





**Fig. 1** Scatter plot of observed versus predicted values for optimal direct mapping approaches. *CLAD* censored least absolute deviation model, *GLM* generalized linear model, *MM* robust MM-estimator

and predicted samples. In the second internal validation procedure (Validation II), the random sample had a sample size of 300, which was generated from the full sample to validate the mapping functions. The final mapping algorithm was developed based on the optimal statistical methods identified from the validation exercises using the full data. All analyses were conducted in Stata version 14.0 (Stata Corp LLC, College Station, TX, USA).

### 2.3.3 Model Comparisons

In this study, the five indicators mentioned above (MAE, MSE, CCC,  $AE > 0.05$  and  $AE > 0.1$ ) were used to evaluate the predictive accuracy of the models. The optimal econometric method for each model specification was identified based on the number of times it has produced the best goodness-of-fit indicators in two types of validation processes (i.e. a total of ten indicators).

## 3 Results

### 3.1 Patient Characteristics

The sociodemographic and clinical characteristics of 607 patients are presented in Table 1. The mean age of participants was 49.0 years (SD 9.8) and 32.8% were either illiterate or completed only primary school education. The majority (88.6%) of participants were married and 50.7% lived in the city. Approximately two-thirds (63.6%) of patients had breast cancer for a duration of up to 3 years; 48% were classified as TNM stage III and IV and approximately half (55%) had a premenopausal status.

### 3.2 Descriptive Statistics and Conceptual Overlap

Table 2 shows the descriptive statistics on the quality of life instruments. The mean utility score derived from the EQ-5D-5L was 0.828 (SD 0.184), and 0.646 (SD 0.125) for the SF-6D. Among all the 23 dimensions of QLQ-BR53, the mean score of the highest dimension (i.e. sexual functioning) was 88.963 (SD 15.933), while the dimension with the lowest score was diarrhea at 10.434 (SD 18.886). The

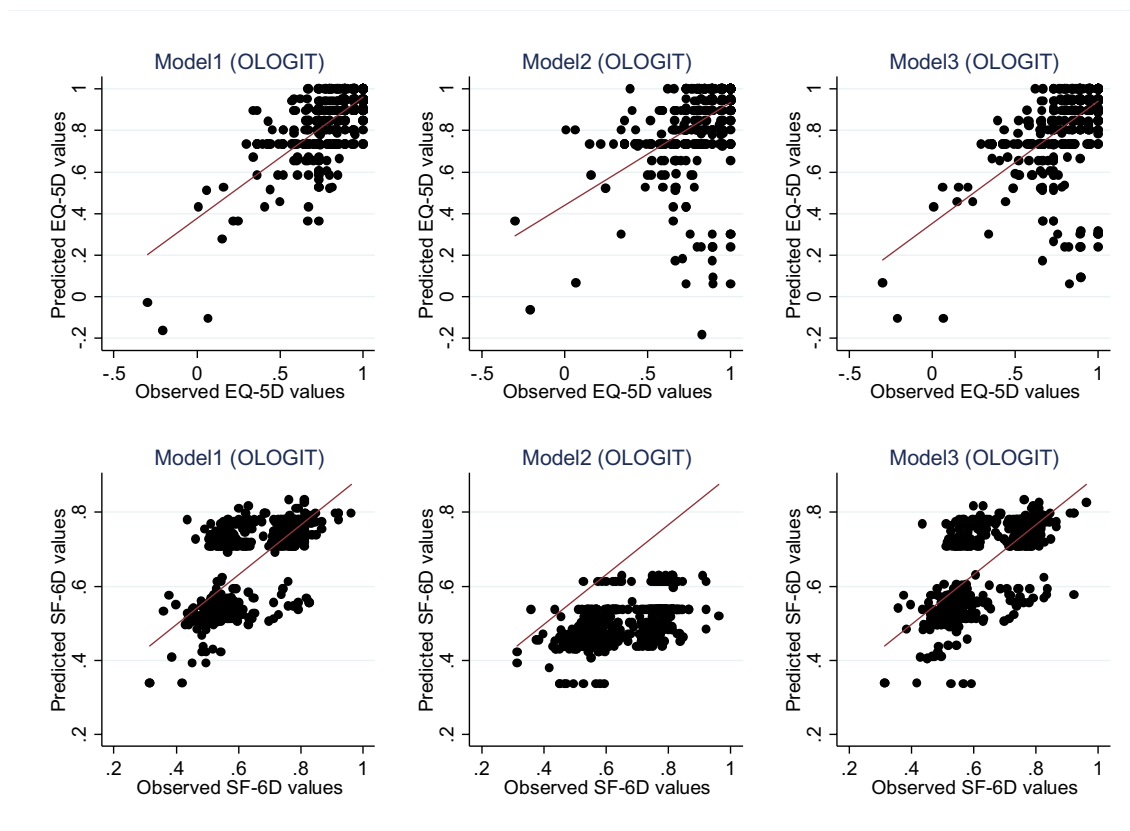


Fig. 2 Scatter plot of observed versus predicted values for the indirect mapping approach. *OLOGIT* ordered logit regression

overlap between QLQ-BR53 and the utility scales (EQ-5D-5L and SF-6D) is presented in Table 3. Among the QLQ-BR53 dimensions, most of the functioning dimensions (e.g. physical functioning and role functioning) and symptom dimensions (e.g. pain and arm symptoms) generally provided stronger ( $r \geq 0.5$ ) correlation with the subscales of EQ-5D-5L and SF-6D. Compared with dimension scores, the utility scores generally provided a strong correlation ( $r \geq 0.5$ , highlighted in bold) with most QLQ-BR53 subscales. All correlations were statistically significant ( $p < 0.01$ ).

### 3.3 Goodness-of-Fit of Mapping Functions

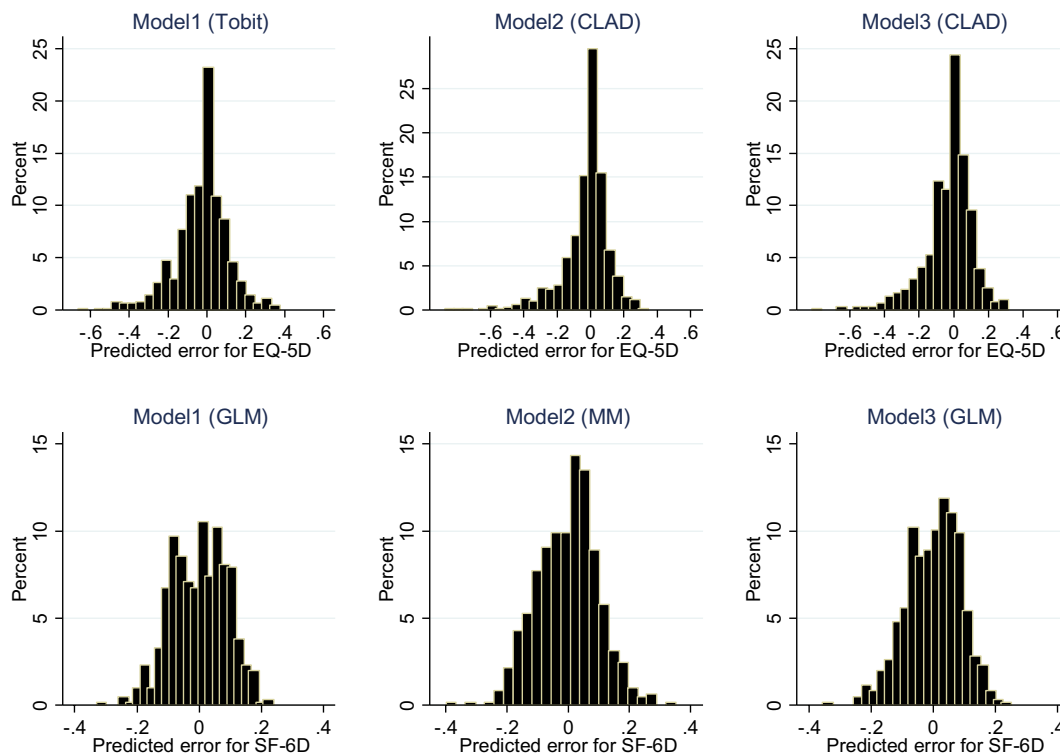
Tables 4 and 5 present a summary of goodness-of-fit statistic results for mapping onto the EQ-5D-5L and SF-6D, based on two validation analyses (i.e. a five-fold pooled validation and a random sample validation), as well as full-sample analyses. It can be seen that the identified optimal econometric methods varied in three model specifications.

When mapping onto the EQ-5D-5L, the Tobit model had the best performance across two types of the validation process for Model 1. The indirect mapping via *OLOGIT* showed good performance, especially in the second validation

process. In Models 2 and 3, the CLAD had the best mapping performance.

Regarding the full-sample results reported in Table 5, the CLAD estimates remain to show the best performance in Models 2 and 3. The performance for Model 1 was mixed, such that the Tobit model had the best performance based on the CCC, while the indirect mapping *OLOGIT* had the best performance on MAE and MSE. However, it should be noted that since no patients reported the fifth level of the anxiety/depression dimension in the EQ-5D-5L, the indirect mapping cannot predict this level. Based on the above considerations, the Tobit model estimates are still chosen as the best mapping function onto the EQ-5D-5L for Model 1, and the CLAD model was chosen for Models 2 and 3. It should also be seen that regardless of the model specifications, all estimators tended to overestimate the lower limit of the EQ-5D-5L utility score. Furthermore, among all estimators, the BETAMIX and *OLOGIT* models were able to predict negative EQ-5D-5L utility scores.

Regarding the performance of mapping onto the SF-6D, the indirect mapping based on *OLOGIT* had the best performance for Models 1 and 3, while the MM-estimator had the best performance for Model 2, in the two validation analyses



**Fig. 3** Predicted error distribution of the optimal direct mapping methods. *CLAD* censored least absolute deviation model, *GLM* generalized linear model, *MM* robust MM-estimator

(Table 4). For Models 1 and 3, except for OLOGIT, the GLM estimates had the best mapping performance among the direct mapping functions. Regarding the full-sample analyses, the above conclusion holds for all three models (Table 5).

Figure 1 shows the scatter plots of observed versus predicted EQ-5D-5L/SF-6D utilities from the optimal econometric methods of each model specification, while Fig. 2 shows the scatter plots of the indirect mapping results based on OLOGIT. The distributions of the prediction errors of direct mapping for the optimal methods of each model specification are shown in Fig. 3, while Fig. 4 shows the predicted error of indirect mapping results.

### 3.4 Optimal Mapping Functions

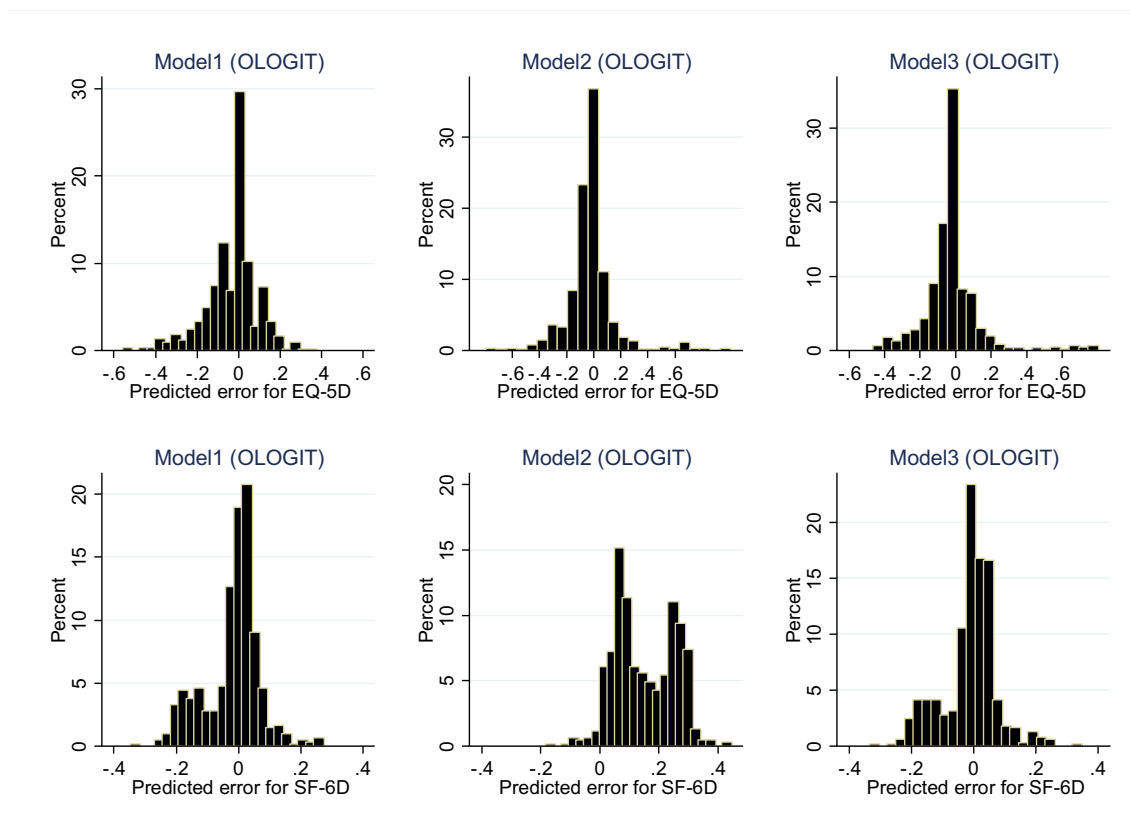
The selected dimension coefficients from the best direct mapping models are reported in Table 6, while the results of the indirect mapping models to predict each dimension are reported in Tables 7, 8, and 9. For OLOGIT regression reported in Tables 7, 8, and 9, it should be noted that the most severe levels of the anxiety/depression dimension in the EQ-5D-5L were not reported by patients in this study

sample. Consequently, this level cannot be predicted accurately based on the results in Tables 7, 8, and 9. Whenever the users have access to both the QLQ-C30 and QLQ-BR23, the mapping algorithms developed based on Model 3 will be preferred, whereas if the user has only included the QLQ-C30 or QLQ-BR23 in the study, mapping algorithms developed from Models 1 or 2 could be used.

To use the mapping functions, the first step is to rescale the raw dimension score onto the 0–1 scale (i.e. raw score divided by 100). Using Model 3 reported in Table 6, the predicted EQ-5D-5L and SF-6D utilities can be calculated as:

$$\begin{aligned} \text{EQ-5D-5L utility} &= 0.6194 + 0.3175 * \text{PF} \\ &+ 0.1300 * \text{EF} - 0.1177 * \text{PA} - 0.3257 * \text{AS} \end{aligned}$$

$$\begin{aligned} \text{SF-6D utility} &= \exp(-1.0632 + 0.2712 \times \text{PF} \\ &+ 0.1608 \times \text{RF} + 0.1302 \times \text{EF} + 0.0739 \times \text{SF} \\ &+ 0.0461 \times \text{CO} + 0.0862 \times \text{QL} + 0.1492 \\ &\times \text{SEF} - 0.0949 \times \text{BS} - 0.1810 \times \text{AS}) \end{aligned}$$



**Fig. 4** Predicted error distribution of the indirect mapping approach. *OLOGIT* ordered logit regression

## 4 Discussion

Accurate measurement and valuation of HRQoL are an important component of economic evaluations of healthcare interventions targeted at breast cancer patients. The QLQ-BR53, EQ-5D-5L and SF-6D have been demonstrated as valid instruments for the measurement of HRQoL in breast cancer patients; however, QLQ-BR53 is not currently preference-based. This study has developed mapping algorithms that can predict EQ-5D-5L and SF-6D utility scores from the QLQ-BR53 to conduct a CUA when the preference-based quality of life instrument is not included. To the best of our knowledge, this is the first mapping study based on the Chinese version of the EORTC QLQ-BR53 for breast cancer patients. There is one existing mapping study that was also developed based on breast cancer patients in China, and this can be used to convert the FACT-B scale onto EQ-5D-5L [38]. Although both FACT-B and QLQ-BR53 are widely used in breast cancer patients, they are two different instruments [39]. The new mapping algorithms reported in this study will further assist researchers to predict health state utility scores based on the breast cancer-specific quality of life instruments in China.

This study differs from previous literature in several ways. First, several mapping algorithms have been developed from the QLQ-C30 onto the EQ-5D-3L, but not yet onto the EQ-5D-5L [13, 40, 41]. This study further extends to include mapping algorithms onto another widely used preference-based instrument, the SF-6D. Without the existence of the tariff from Mainland China, the Hong Kong tariff, which has been widely used among Chinese populations, was adopted. Based on the reported indirect mapping functions, the results reported in this study can be used when the Chinese-specific tariff of Mainland China is available in the future. For the EQ-5D-5L and SF-6D studied here, mapping onto SF-6D tends to have a better performance than the EQ-5D-5L in general (based on lower predicted errors and higher CCC values).

Second, for breast cancer clinical trials, if two instruments are included, the study found that using dimensions from both instruments can lead to a better prediction [13]. Comparing our study with the Korean study by Kim et al. [13], similar model specifications were used and some of the final predictors were the same, e.g. the physical functioning dimension and the symptom dimensions of pain, dyspnea, systemic therapy side effects, and arm symptoms (all of which were statistically significant at the 0.05 level).

**Table 6** Direct mapping equations from the QLQ-BR53 index to health state utilities

	EQ-5D-5L	SF-6D
<i>Model 1</i>	Tobit	GLM
PF	0.3154 (0.0598)	0.3269 (0.0524)
RF	0.1167 (0.0418)	0.2316 (0.0356)
EF	0.1337 (0.0494)	0.1555 (0.0392)
SF	0.1053 (0.0360)	0.1129 (0.0292)
NV		-0.0782 (0.0373)
PA	-0.2187 (0.0485)	
DY	-0.1045 (0.0391)	
CO		0.0570 (0.0240)
Age	0.0312 (0.0145)	
Constant	0.4011 (0.0591)	-1.0675 (0.0433)
<i>Model 2</i>	CLAD	MM
BI	0.0925 (0.0263)	
SEF		0.1458 (0.0270)
ST		-0.1867 (0.0363)
BS	-0.2147 (0.0535)	-0.0921 (0.0356)
AS	-0.3768 (0.0519)	-0.2664 (0.0432)
Constant	0.9075 (0.0263)	0.6352 (0.0275)
<i>Model 3</i>	CLAD	GLM
PF	0.3175 (0.0633)	0.2712 (0.0509)
RF		0.1608 (0.0352)
EF	0.1300 (0.0331)	0.1302 (0.0371)
SF		0.0739 (0.0280)
PA	-0.1177 (0.0384)	
CO		0.0461 (0.0222)
QL		0.0862 (0.0409)
SEF		0.1492 (0.0366)
BS		-0.0949 (0.0444)
AS	-0.3257 (0.0430)	-0.1810 (0.0434)
Constant	0.6194 (0.0538)	-1.0632 (0.0599)

Standard errors are shown in parentheses and were significant at 5% [ $N=607$ ]

CLAD censored least absolute deviation, GLM generalized linear model, MM robust MM-estimator, PF physical functioning, RF role functioning, EF emotional functioning, SF social functioning, NV nausea and vomiting, PA pain, DY dyspnea, CO constipation, QL global health status, BI body image, SEF sexual functioning, ST systemic therapy side effects, BS breast symptoms, AS arm symptoms

To some extent, this similarity reflects a robust relationship between some QLQ-BR53 dimensions and EQ-5D utility. From the goodness-of-fit results, we can see that the mapping performance based on Model 3 was always better than Models 1 or 2; this finding suggests that in predicting health state utilities of breast cancer patients, the use of QLQ-BR53 will lead to a better prediction accuracy than using either QLQ-C30 or QLQ-BR23 alone.

Third, previous mapping studies have only used the OLS method to derive a mapping algorithm in this context [13].

Although the OLS is a popular method, its prediction performance may not be as good as other statistical methods. As demonstrated in this study, the OLS estimates have not been selected as the optimal algorithms. It should be noted that owing to the different versions of the EQ-5D instruments (i.e. 3-level vs 5-level EQ-5D used in this paper), as well as the different country-specific tariffs used, it is also impossible to directly compare the mapping performance between this study and previous literature [13, 40]. On the other hand, similar to previous mapping studies, almost all mapping algorithms tend to overestimate the utilities for patients in poor health [42].

There are some limitations to this study. First, only breast cancer patients were used in the study, therefore it is unclear to what extent the mapping algorithms from QLQ-C30 onto EQ-5D-5L/SF-6D can be more widely used for other cancer patients. Second, although the predicted mean health state utility scores at the sample level are very close to the observed mean utility, as commonly reported in the literature, the mapping performance at the lower end of the utility distribution is not good. For example, among the econometric methods used, only BETAMIX and OLOGIT can predict the negative EQ-5D-5L utility scores; however, based on a wide range of goodness-of-fit indicators, the overall mapping performance of these two methods may not be as good as other methods. Caution is therefore warranted when applying mapping algorithms to patients in poor health status. Third, the developed mapping algorithms should be externally validated. In particular, external validation using a longitudinal dataset will be helpful to explore to what extent the incremental utility can be accurately calculated based on the mapping algorithms.

Finally, mapping algorithms should serve as a second-best solution to generate health utilities from the non-preference-based, disease-specific quality of life instruments [23]. Recently, King et al. [43] developed a health state classification system from the EORTC QLQ-C30, called the EORTC Quality of Life Utility Measure-Core 10 dimensions (QLU-C10D). The responses to the QLQ-C30 can be converted into QLU-C10D utility scores for conducting cost-utility analyses. However, the QLU-C10D only contains information from the QLQ-C30, and, as such, more breast cancer-specific dimensions that are captured using the QLQ-BR23 will be omitted.

## 5 Conclusions

This study reported mapping algorithms from the QLQ-C30 and/or QLQ-BR23 onto EQ-5D-5L or SF-6D utilities based on breast cancer patients in China. Outputs from this study can be used in CUAs to prioritise health resources. Further studies are warranted to externally validate mapping

Table 7 Indirect mapping equations from the QLQ-BR53 index to each of the EQ-5D-5L and SF-6D dimensions (ordered logit regression): Model 1

	EQ-5D-5L dimensions					SF-6D dimensions					
	Mobility	Self-care	Usual activities	Pain/discomfort	Anxiety/depression	Physical health	Role limitation	Social functioning	Pain	Mental health	Vitality
PF	-4.682244 (0.829)	-3.574644 (0.758)	-2.808575 (0.772)	-2.466179 (0.628)	-4.187416 (0.708)	-6.59834 (0.811)	-3.80190 (0.812)	-3.12167 (0.704)	-3.12167 (0.704)	-2.01932 (0.644)	
RF	-1.69769 (0.567)	-2.728378 (0.541)	-1.825835 (0.515)		1.66452 (0.778)	-1.78072 (0.479)	-2.31575 (0.506)	-1.917918 (0.366)	-2.30671 (0.513)		-1.69410 (0.454)
EF					-4.187416 (0.708)	-2.16664 (0.542)				-1.98262 (0.584)	
CF			-1.378187 (0.667)		1.66452 (0.778)		-2.50256 (0.620)				
SF		-1.414439 (0.477)	-1.481716 (0.502)	-1.271212 (0.402)	-1.166634 (0.442)				-1.76259 (0.423)		-1.311703 (0.417)
FA				1.689312 (0.689)						3.39785 (0.743)	
NV	1.554053 (0.608)		1.044594 (0.522)	3.094723 (0.612)	2.828986 (0.603)		1.440056 (0.541)		2.85453 (0.563)	-1.5244 (0.613)	1.068812 (0.548)
PA	1.53265 (0.675)	1.244637 (0.600)			1.358326 (0.443)						
DY			1.70187 (0.519)		-1.094406 (0.351)						
CO							-1.45114 (0.376)	-1.088579 (0.320)	-0.95445 (0.326)	-0.7985 (0.339)	
DI	1.766741 (0.577)	1.60608 (0.498)				1.111281 (0.523)					2.16676 (0.499)
FI								1.910503 (0.326)			
QL								-1.579121 (0.579)			-2.68850 (0.622)
Age	-0.371662 (0.222)				-0.394128 (0.177)	-0.37404 (0.165)					
/cut1	-3.512304 (0.776)	-4.912437 (0.689)	-4.912635 (0.672)	-2.001745 (0.635)	-2.905279 (0.695)	-13.7167 (0.808)	-7.13815 (0.827)	-4.254294 (0.548)	-7.22993 (0.661)	-6.72875 (0.822)	-6.95818 (0.637)
/cut2	0.049041 (0.764)	-1.718622 (0.653)	-1.712363 (0.641)	0.901141 (0.626)	0.165227 (0.686)	-10.0595 (0.705)	-5.73789 (0.807)	-1.512123 (0.526)	-6.83026 (0.658)	-3.03884 (0.767)	-4.549627 (0.588)
/cut3	2.29953 (0.868)	0.680027 (0.716)	0.673185 (0.673)	3.88286 (0.718)	3.429414 (0.816)	-8.07790 (0.664)	-5.44282 (0.803)	1.393067 (0.526)	-3.55773 (0.614)	-0.03478 (0.756)	-1.232826 (0.559)
/cut4	3.736998 (1.062)	1.671403 (0.789)	2.038221 (0.765)	6.085482 (0.973)		-7.73075 (0.656)		2.539772 (0.550)	-1.13883 (0.598)	1.72631 (0.797)	0.892137 (0.580)
/cut5						-3.28309 (0.566)			1.66276 (0.709)		

Standard errors are shown in parentheses and were significant at 5% [N=607]

PF physical functioning, RF role functioning, EF emotional functioning, CF cognitive functioning, SF social functioning, FA fatigue, NV nausea and vomiting, PA pain, DY dyspnea, CO constipation, DI diarrhea, FI financial difficulties, QL global health status

**Table 8** Indirect mapping equations from the QLQ-BR53 index to each of the EQ-5D-5L and SF-6D dimensions (ordered logit regression): Model 2

	EQ-5D-5L dimensions					SF-6D dimensions					
	Mobility	Self-care	Usual activities	Pain/discomfort	Anxiety/depression	Physical health	Role limitation	Social functioning	Pain	Mental health	Vitality
BI	-1.770685 (0.524)			-0.940851 (0.422)	-1.470641 (0.448)		-1.616481 (0.448)			-1.113137 (0.347)	
SEF	-3.67263 (0.632)	-2.447719 (0.606)	-1.978725 (0.600)			-2.749243 (0.538)	-1.893596 (0.571)	-2.585856 (0.842)	-2.09115 (0.785)	0.7621142 (0.451)	
SEE								2.245754 (0.702)	1.367306 (0.656)		
FU								-0.742851 (0.268)			
ST	3.282659 (0.869)	3.898364 (0.876)	4.517662 (0.878)	2.233012 (0.687)		2.230921 (0.599)			2.457173 (0.595)		3.579729 (0.602)
BS	1.853166 (0.713)	1.988388 (0.661)	1.443024 (0.655)		1.470418 (0.628)	3.09838 (0.645)	1.631904 (0.644)			2.690664 (0.468)	
AS	2.324685 (0.775)	4.016127 (0.716)	4.318164 (0.705)	4.734758 (0.545)	3.244271 (0.677)	1.355898 (0.630)	3.525648 (0.665)	2.448327 (0.467)	4.692161 (0.552)		1.860215 (0.517)
HL					1.167287 (0.359)						
Age	-0.573442 (0.222)				-0.459931 (0.179)	-0.437322 (0.161)		0.4105074 (0.163)			-0.214779 (0.165)
/cut1	-2.534322 (0.818)	0.1858316 (0.569)	0.491859 (0.569)	0.173504 (0.396)	-0.762417 (0.468)	-6.612399 (0.669)	-2.309402 (0.641)	-2.288548 (0.564)	-1.265425 (0.513)	-3.586452 (0.568)	-2.657972 (0.337)
/cut2	0.718400 (0.814)	3.240103 (0.590)	3.541397 (0.593)	3.0923 (0.423)	2.432149 (0.484)	-3.125303 (0.569)	-0.9383925 (0.636)	0.3617989 (0.547)	-0.887304 (0.511)	0.035522 (0.502)	-0.359328 (0.273)
/cut3	2.580924 (0.914)	5.523344 (0.717)	5.688949 (0.697)	5.76641 (0.526)	5.503377 (0.664)	-1.392833 (0.557)	-0.6671108 (0.635)	3.113689 (0.569)	2.085726 (0.515)	2.962511 (0.523)	2.720921 (0.303)
/cut4	3.792573 (1.097)	6.48934 (0.835)	7.133069 (0.892)	7.717068 (0.850)		-1.097505 (0.556)		4.230341 (0.611)	4.07562 (0.544)	4.699523 (0.597)	4.614117 (0.369)
/cut5						2.441905 (0.586)			6.130516 (0.669)		

Standard errors are shown in parentheses and were significant at 5% [ $N = 607$ ]

BI body image, SEF sexual functioning, SEE sexual enjoyment, FU future perspective, ST systemic therapy side effects, BS breast symptoms, AS arm symptoms, HL upset by hair loss



**Table 9** Indirect mapping equations from the QLQ-BR53 index to each of the EQ-5D-5L and SF-6D dimensions (ordered logit regression): Model 3

	EQ-5D-5L dimensions				SF-6D dimensions						
	Mobility	Self-care	Usual activities	Pain/discomfort	Anxiety/depression	Physical health	Role limitation	Social functioning	Pain	Mental health	Vitality
PF	-4.303375 (0.855)	-3.793802 (0.833)	-3.040293 (0.848)	-2.235587 (0.612)		-6.014033 (0.825)	-3.372381 (0.832)		-2.886249 (0.726)	-1.707856 (0.705)	
RF	-1.356188 (0.586)	-2.527803 (0.588)	-1.510968 (0.572)			-1.503527 (0.506)	-1.110001 (0.549)	-1.743204 (0.472)	-1.900828 (0.533)		-1.510061 (0.460)
EF					-3.672195 (0.729)	-2.142621 (0.573)				-1.41117 (0.600)	
CF			-1.414393 (0.707)		1.838307 (0.764)		-1.86926 (0.655)				
SF								-1.873297 (0.618)	-1.42661 (0.435)		
FA								-1.577115 (0.643)	3.13126 (0.754)		
NV							1.287157 (0.581)				
PA					1.810029 (0.660)				2.49058 (0.611)	-2.179504 (0.649)	
DY			1.457423 (0.538)								
SL		0.999319 (0.479)									
AP											
CO		-1.127334 (0.480)			-0.826955 (0.366)		-1.054877 (0.396)	-0.893840 (0.343)	-0.732079 (0.338)	-0.908480 (0.342)	
DI	1.372337 (0.566)					0.822431 (0.405)					1.992696 (0.509)
FI								1.120471 (0.482)			1.045942 (0.340)
QL							1.369816 (0.688)	-2.222406 (0.701)			-2.645619 (0.625)
BI				-1.429448 (0.471)		-0.894863 (0.381)					
SEF	-2.693691 (0.676)		-1.697733 (0.679)			-2.206971 (0.569)	-2.790863 (0.909)		-1.977444 (0.827)		
SEE							2.361663 (0.757)		1.774086 (0.700)		
ST	2.852405 (0.825)	2.704276 (0.980)	2.582008 (0.944)								

Table 9 (continued)

	EQ-5D-5L dimensions					SF-6D dimensions					
	Mobility	Self-care	Usual activities	Pain/discomfort	Anxiety/depression	Physical health	Role limitation	Social functioning	Pain	Mental health	Vitality
BS						1.391942 (0.563)				1.924606 (0.545)	
AS	1.965393 (0.689)	2.993068 (0.763)	3.45946 (0.651)	3.403158 (0.577)	3.20544 (0.599)		3.075074 (0.599)	1.292145 (0.573)	1.960109 (0.596)		
HL		1.171744 (0.461)			1.201704 (0.338)		0.8396433 (0.333)	-0.731945 (0.329)			
Age					-0.480983 (0.185)		-0.357182 (0.167)				
/cut1	-4.411157 (0.907)	-3.788261 (0.883)	-5.309539 (1.186)	-1.627036 (0.602)	-1.903266 (0.821)	-14.17312 (1.014)	-4.516575 (1.076)	-6.135442 (1.084)	-6.381111 (0.795)	-5.15641 (0.880)	-5.591613 (0.714)
/cut2	-0.817912 (0.878)	-0.342988 (0.862)	-1.808992 (1.158)	1.439345 (0.600)	1.429504 (0.823)	-10.49517 (0.935)	-2.998156 (1.066)	-3.201487 (1.055)	-5.975047 (0.792)	-1.435427 (0.835)	-3.163679 (0.676)
/cut3	1.287816 (0.978)	2.316632 (0.909)	0.659952 (1.182)	4.3368 (0.675)	4.742255 (0.931)	-8.414456 (0.900)	-2.697508 (1.064)	-0.109285 (1.053)	-2.617623 (0.758)	1.667626 (0.837)	0.181452 (0.667)
/cut4	2.678323 (1.167)	3.63757 (0.993)	2.498703 (1.287)	6.629782 (0.960)		-8.038223 (0.893)		1.091485 (1.073)	-0.140925 (0.754)	3.462307 (0.883)	2.305964 (0.687)
/cut5						-3.566602 (0.826)			2.507444 (0.862)		

Standard errors are shown in parentheses and were significant at 5% [N = 607]

PF physical functioning, RF role functioning, EF emotional functioning, CF cognitive functioning, FA fatigue, MV nausea and vomiting, PA pain, DY dyspnea, SL insomnia, AP appetite loss, CO constipation, DI diarrhea, FI financial difficulties, QL global health status, BI body image, SEF sexual functioning, SEE sexual enjoyment, ST systemic therapy side effects, BS breast symptoms AS arm symptoms, HL upset by hair loss

algorithms and explore the use of these mapping algorithms in other cancer patients.

**Acknowledgements** The authors thank all participants for their time and effort.

**Author Contribution** SL and QS conceived and designed this study; MW participated in the acquisition of the data; TL analyzed the data, interpreted the results and wrote the first draft of the manuscript; and GC supervised the statistical analysis, gave feedback on the manuscript and revised it critically for important intellectual content. All authors have read and approved the final version of the manuscript.

**Data Availability Statement** The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

## Compliance with Ethical Standards

**Funding** Financial support was received from Shandong Provincial Natural Science Foundation, China (ZR2013GM023).

**Conflict of interest** Tong Liu, Shunping Li, Min Wang, Qiang Sun, and Gang Chen declare no conflicts of interest.

**Ethical approval** Ethical approval (reference no. 20131002) was obtained from the Ethics Review Board of the School of Public Health, Shandong University. This research adhered to the tenets of the Declaration of Helsinki.

**Informed consent** Informed consent was obtained from all individual participants included in the studies.

## References

- Key TJ, Verkasalo PK, Banks E. Epidemiology of breast cancer. *Lancet Oncol*. 2001;2(3):133–40.
- Ginsburg O, Bray F, Coleman MP, Vanderpuye V, Eniu A, Kotha SR, et al. The global burden of women's cancers: a grand challenge in global health. *Lancet*. 2017;389(10071):847–60. [https://doi.org/10.1016/s0140-6736\(16\)31392-7](https://doi.org/10.1016/s0140-6736(16)31392-7).
- Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Pineros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*. 2019;144(8):1941–53. <https://doi.org/10.1002/ijc.31937>.
- Chen WQ, Zheng RS, Baade PD, Zhang SW, Zeng HM, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin*. 2016;66(2):115–32. <https://doi.org/10.3322/caac.21338>.
- Earle CC, Chapman RH, Baker CS, Bell CM, Stone PW, Sandberg EA, et al. Systematic overview of cost-utility assessments in oncology. *J Clin Oncol*. 2000;18(18):3302–17. <https://doi.org/10.1200/jco.2000.18.18.3302>.
- Wan C, Tang X, Tu XM, Feng C, Messing S, Meng Q, et al. Psychometric properties of the simplified Chinese version of the EORTC QLQ-BR53 for measuring quality of life for breast cancer patients. *Breast Cancer Res Treat*. 2007;105(2):187–93. <https://doi.org/10.1007/s10549-006-9443-1>.
- Zhang Z, Zhang X, Wei L, Lin Y, Wu D, Xie S, et al. Questionnaire to assess quality of life in patients with breast cancer—validation of the Chinese version of the EORTC QLQ-BR 53. *Breast*. 2017;32:87–92. <https://doi.org/10.1016/j.breast.2016.12.019>.
- Liang X-F, Ma D-C, Ding Z-Y, Liu Z-Z, Guo F, Liu L, et al. Autologous cytokine-induced killer cells therapy on the quality of life of patients with breast cancer after adjuvant chemotherapy: a prospective study [in Chinese]. *Zhonghua zhong liu za zhi*. 2013;35(10):764–8.
- Wailoo AJ, Hernandez-Alava M, Manca A, Mejia A, Ray J, Crawford B, et al. Mapping to estimate health-state utility from non-preference-based outcome measures: an ISPOR good practices for outcomes research task force report. *Value Health*. 2017;20(1):18–27. <https://doi.org/10.1016/j.jval.2016.11.006>.
- Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ*. 2010;11(2):215–25. <https://doi.org/10.1007/s10198-009-0168-z>.
- Dakin H. Review of studies mapping from quality of life or clinical measures to EQ-5D: an online database. *Health Qual Life Outcomes*. 2013;11:151. <https://doi.org/10.1186/1477-7525-11-151>.
- Kang H, Ko SK, Kim EJ. Mapping the cancer-specific eortc QLQ-c30 and eortc QLQ-br23 to the generic EQ-5d in metastatic breast cancer patients. *Value Health*. 2011;14(7):A458. <https://doi.org/10.1016/j.jval.2011.08.1230>.
- Kim EJ, Ko SK, Kang HY. Mapping the cancer-specific EORTC QLQ-C30 and EORTC QLQ-BR23 to the generic EQ-5D in metastatic breast cancer patients. *Qual Life Res*. 2012;21(7):1193–203. <https://doi.org/10.1007/s11136-011-0037-y>.
- Fayers P, Aaronson N, Bjordal K, Groenvold M, Curran D, Bottomley A. The EORTC QLQ-C30 scoring manual. 2001. Brussels: European Organisation for Research and Treatment of Cancer; 2015. p. 3.
- Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727–36. <https://doi.org/10.1007/s11136-011-9903-x>.
- Luo N, Liu G, Li MH, Guan HJ, Jin XJ, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. *Value Health*. 2017;20(4):662–9. <https://doi.org/10.1016/j.jval.2016.11.016>.
- Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. *J Clin Epidemiol*. 1998;51(11):1115–28. [https://doi.org/10.1016/s0895-4356\(98\)00103-6](https://doi.org/10.1016/s0895-4356(98)00103-6).
- Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002;21(2):271–92. [https://doi.org/10.1016/s0167-6296\(01\)00130-8](https://doi.org/10.1016/s0167-6296(01)00130-8).
- Lam C, Brazier J, Mcghee S. Valuation of the SF-6D health states is feasible, acceptable, reliable, and valid in a Chinese population. *Value Health*. 2008;11(2):295–303.
- Petrou S, Rivero-Arias O, Dakin H, Longworth L, Oppe M, Froud R, et al. The MAPS reporting statement for studies mapping onto generic preference-based outcome measures: explanation and elaboration. *Pharmacoeconomics*. 2015;33(10):993–1011. <https://doi.org/10.1007/s40273-015-0312-9>.
- Longworth L, Yang Y, Young T, Mulhern B, Hernandez Alava M, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess*. 2014;18(9):1–224. <https://doi.org/10.3310/hta18090>.
- Dakin H, Abel L, Burns R, Yang YL. Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: an online database and application of the MAPS statement. *Health Qual Life Outcomes*. 2018;16(1):31. <https://doi.org/10.1186/s12955-018-0857-3>.
- Longworth L, Rowen D. Mapping to obtain EQ-5D utility values for use in NICE health technology assessments. *Value Health*. 2013;16(1):202–10. <https://doi.org/10.1016/j.jval.2012.10.010>.

24. Kontodimopoulos N. The potential for a generally applicable mapping model between QLQ-C30 and SF-6D in patients with different cancers: a comparison of regression-based methods. *Qual Life Res.* 2015;24(6):1535–44. <https://doi.org/10.1007/s11136-014-0857-7>.
25. Whitehurst DG, Bryan S. Another study showing that two preference-based measures of health-related quality of life (EQ-5D and SF-6D) are not interchangeable. But why should we expect them to be? *Value Health.* 2011;14(4):531–8.
26. Powell JL. Least absolute deviations estimation for the censored regression model. *J Econometr.* 1984;25(3):303–25.
27. McCulloch CE. Generalized linear models. *J Am Stat Assoc.* 2000;95(452):1320–4. <https://doi.org/10.2307/2669780>.
28. Yohai VJ. High breakdown-point and high efficiency robust estimates for regression. *Ann Stat.* 1987;15(2):642–56.
29. Chen G, Stevens K, Rowen D, Ratcliffe J. From KIDSCREEN-10 to CHU9D: creating a unique mapping algorithm for application in economic evaluation. *Health Qual Life Outcomes.* 2014;12:134. <https://doi.org/10.1186/s12955-014-0134-z>.
30. Chen G, Khan MA, Iezzi A, Ratcliffe J, Richardson J. Mapping between 6 multiattribute utility instruments. *Med Decis Making.* 2016;36(2):160–75. <https://doi.org/10.1177/0272989x15578127>.
31. Gray LA, Alava MH. A command for fitting mixture regression models for bounded dependent variables using the beta distribution. *Stata J.* 2018;18(1):51–755.
32. Ospina R, Ferrari SL. A general class of zero-or-one inflated beta regression models. *Comput Stat Data Anal.* 2012;56(6):1609–23.
33. Zheng Y, Tang K, Ye L, Ai Z, Wu B. Mapping the neck disability index to SF-6D in patients with chronic neck pain. *Health Qual Life Outcomes.* 2016;14:21. <https://doi.org/10.1186/s12955-016-0422-x>.
34. Kaambwa B, Ratcliffe J. Predicting EuroQoL 5 dimensions 5 levels (EQ-5D-5L) utilities from older people's quality of life brief questionnaire (OPQoL-Brief) scores. *Patient.* 2018;11(1):39–54. <https://doi.org/10.1007/s40271-017-0259-3>.
35. Hernandez Alava M, Wailoo A, Wolfe F, Michaud K. A Comparison of direct and indirect methods for the estimation of health utilities from clinical outcomes. *Med Decis Making.* 2014;34(7):919–30.
36. Chen G, Garcia-Gordillo MA, Collado-Mateo D, Del Pozo-Cruz B, Adsuar JC, Cordero-Ferrera JM, et al. Converting Parkinson-specific scores into health state utilities to assess cost-utility analysis. *Patient.* 2018;11(6):665–75. <https://doi.org/10.1007/s40271-018-0317-5>.
37. Long JS. *Regression models for categorical and limited dependent variables.* Sage Publications; 1997.
38. Yang Q, Yu XX, Zhang W, Li H. Mapping function from FACT-B to EQ-5D-5 L using multiple modelling approaches: data from breast cancer patients in China. *Health Qual Life Outcomes.* 2019. <https://doi.org/10.1186/s12955-019-1224-8>.
39. Kemmler G, Holzner B, Kopp M, Dunser M, Margreiter R, Greil R, et al. Comparison of two quality-of-life instruments for cancer patients: the functional assessment of cancer therapy-general and the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30. *J Clin Oncol.* 1999;17(9):2932–40. <https://doi.org/10.1200/jco.1999.17.9.2932>.
40. Crott R, Briggs A. Mapping the QLQ-C30 quality of life cancer questionnaire to EQ-5D patient preferences. *Eur J Health Econ.* 2010;11(4):427–34.
41. McKenzie L, Pol MVD. Mapping the EORTC QLQ C-30 onto the EQ-5D instrument: the potential to estimate qalys without generic preference data. *Value Health.* 2009;12(1):167–71.
42. Doble B, Lorgelly P. Mapping the EORTC QLQ-C30 onto the EQ-5D-3L: assessing the external validity of existing mapping algorithms. *Qual Life Res.* 2016;25(4):891–911. <https://doi.org/10.1007/s11136-015-1116-2>.
43. King MT, Costa DSJ, Aaronson NK, Brazier JE, Cella DF, Fayers PM, et al. QLU-C10D: a health state classification system for a multi-attribute utility measure based on the EORTC QLQ-C30. *Qual Life Res.* 2016;25(3):625–36. <https://doi.org/10.1007/s11136-015-1217-y>.