ScienceDirect
Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Preference-Based Assessments

Investigating 5-Level EQ-5D (EQ-5D-5L) Values Based on Preferences of Patients With Heart Disease

Mihir Gandhi, PhD, Ru San Tan, MBBS, Shir Lynn Lim, MBBS, Kim Rand, PhD, Carolyn S.P. Lam, PhD, Nan Luo, PhD, Yin Bun Cheung, PhD

ABSTRACT

Objectives: Several studies have shown that patients with heart disease value hypothetical health states differently from the general population. We aimed to investigate the health preferences of patients with heart disease and develop a value set for the 5-level EQ-5D (EQ-5D-5L) based on these patient preferences.

Methods: Patients with confirmed heart disease were recruited from 2 hospitals in Singapore. A total of 86 EQ-5D-5L health states (10 per patient) were valued using a composite time trade-off method according to the international valuation protocol for EQ-5D-5L; 20-parameter linear models and 8-parameter cross-attribute level effects models with and without an N45 term (indicating whether any health state dimension at level 4 or 5 existed) were estimated. Each model included patient-specific random intercepts. Model performance was evaluated for out-of-sample and in-sample predictive accuracy in terms of root mean square error. The discriminative ability of the utility values was assessed using heart disease-related functional classes.

Results: A total of 576 patients were included in the analysis. The preferred model, with the lowest out-of-sample root mean square error, was a 20-parameter linear model including N45. Predicted utility values ranged from −0.727 for the worst state to 1 for full health; the value for the second-best state was 0.981. Utility values demonstrated good discriminative ability in differentiating among patients of varied functional classes.

Conclusions: An EQ-5D-5L value set representing the preferences of patients with heart disease was developed. The value set could be used for patient-centric economic evaluation and health-related quality of life assessment for patients with heart disease.

Keywords: cardiovascular disease, EQ-5D, EQ-5D-5L, heart disease, patient preference, preference, utility, value set.

VALUE HEALTH. 2022; 25(3):451–460

Introduction

Medical costs are escalating with population aging and advances in healthcare technologies,[1] and these changes are placing pressure on national healthcare budgets. Health technology assessment (HTA) helps make efficient use of healthcare budgets. Using quality-adjusted life-years, HTA evaluates the costs of new treatment taking into consideration survival benefits and effects on health-related quality of life (HRQoL). Quality-adjusted life-years are typically obtained from a generic preference-based instrument such as the EQ-5D that provides a utility value that is multiplied by the duration lived in a health state. Utility values are usually estimated by asking people to assign values to specific hypothetical health states that vary in severity from mild to extremely severe.

HTA methods and processes have been criticized for not being sufficiently "patient-centric."[2] For example, several countries recommend that the reference case analysis be based on a societal perspective.[3] Nevertheless, there are doubts about whether members of the general public who are relatively healthy can appreciate the health states that they are being asked to value. Therefore, there are arguments that support using the preferences of patients who have experienced health states of varying severity.[4] A patient is could be an individual who is currently experiencing the health state being valued or might have experienced it in the past or have experienced other health states similar to that being valued.[5] The Dental and Pharmaceutical Benefits Agency in Sweden recommends that the preferences of persons who have experienced the particular health condition that is being assessed be used in economic evaluations.[6] For medical technology evaluations, the Agency for Care Effectiveness in Singapore recommends the use of preferences based both on patients with the condition and on the general public.[7,8]

The choice of using patient or general public preferences depends on the purpose and context of the evaluation. General public values are desirable when the values are used to inform decisions that allocate societal resources, whereas patient values may be more appropriate when making treatment decisions guided by patient health preferences. Patient preferences are essential for patient-centric healthcare decisions. In many counties, patient preferences are also crucial for economic evaluations, because patients themselves bear most healthcare costs.[9]

Empirical studies have shown systematic differences in the valuing of hypothetical health states by the general public and by patients with certain health conditions, such as heart disease,[10-12] that are not explained by differences in sociodemographic characteristics. Pickard et al[12] and Gandhi et al[11] showed that patients with heart disease give higher values than the general public for the 3-level EQ-5D (EQ-5D-3L) health states. Differences in values given by patients with heart disease and healthy people were also reported for the 5-level EQ-5D (EQ-5D-5L) health states,[10] along with the impact of these differences on utility gain estimates. These differences could occur for several reasons: variation in life experiences, uncertainty about life, adaptation to suboptimal health conditions, and healthcare costs. These findings support the use of utility values based on patient preferences for patient-centric healthcare decision making.

In addition to economic evaluation, EQ-5D-5L is a widely used instrument for evaluating the impact of healthcare interventions on HRQoL. A summary score based on EQ-5D-5L dimensions that reflects patient perspective is useful for clinical studies. Nevertheless, currently no such score is available for EQ-5D-5L. Given that different patient populations may weigh the dimensions and their severity levels differently, a score based on preferences of patients with heart disease would be the best for use in this patient population.

Cardiovascular diseases, which include ischemic heart disease, stroke, peripheral arterial disease, heart failure, and several other cardiac and vascular conditions, contribute to >400 million new cases, 18 million deaths (31% of all mortality), and 36 million years of lived-with-disability per year worldwide.[13] Considering the disease prevalence, burden, and health preferences, a utility value set based on preferences of patients with heart disease will potentially have a significant impact on the evaluation of emerging therapies for cardiovascular disease.

In this study, we aimed to investigate the preferences of patients with heart disease for health states defined by the EQ-5D-5L descriptive system and develop a utility value set based on these patient preferences. We defined a patient as an individual who currently has a heart disease and previous experience of hospitalization because of a heart problem. A patient is expected to have past experience of health states similar to that to be valued. The EQ-5D-5L is a new version of the widely used EQ-5D-3L and has demonstrated better measurement properties than the previous version.[14]

Methods

Study Design and Participants

This was a cross-sectional study involving face-to-face interviews of patients with heart disease who were receiving treatment at the 2 largest cardiovascular tertiary hospitals in Singapore, a multiethnic Asian city-state. Consecutive patients were approached during their regular outpatient clinic visits.

The eligibility criteria for the study were (1) adult patient (21 years or older) with one or more types of clinically confirmed heart disease (ischemic heart disease, heart failure, heart rhythm disorder, valvular heart disease) and previous hospitalization for heart disease-related conditions; (2) physically and mentally well enough to participate in a 30-minute interview; and (3) able to read and communicate in either English or Chinese. The eligibility criterion of previous hospitalization was included to ensure that all of the study participants had experienced a severe health state. The Singapore resident population constitutes 74% Chinese, 13% Malay, 9% Indian, and 3% others. More than 85% of Native Americans and Malay are literate in English.[15] Hence, the eligibility criteria for language cover all 3 major ethnic populations in Singapore. The diagnosis of heart disease was based on internationally accepted criteria as applied by the participants' managing cardiologists.

An informed consent was obtained from all participants. The study was approved by the ethics boards of the respective hospitals.

Valuation Interview

The participants were interviewed in quiet areas in clinics. Each participant was interviewed by a trained interviewer in the language of their preference. The interviewer team comprised 4 bilingual interviewers who could fluently read and speak English and Chinese and had previous experience in conducting patient interviews. All the interviewers were trained in the valuation protocol and had conducted at least 5 practice interviews.

Each interview comprised 2 parts: the first involved self-administration of paper forms, and the second involved interviewer-guided computer-based valuation tasks. In the first part, participants self-reported their sociodemographic information and health profiles using the EQ-5D-5L along with a visual analog scale (EQ-VAS) and the HeartQoL (heart disease-specific HRQoL instrument); they also reported their functional status using the New York Heart Association (NYHA) classification and the Canadian Cardiovascular Society (CCS) classification of angina. In the second part, participants valued EQ-5D-5L health states using the composite time trade-off (cTTO) module of the EuroQol Portable Valuation Technology software version 1.7 running from a laptop. The interviewers followed a standard script in all interviews. In a previous study, the script and valuation tasks administered using very similar software were tested and shown to be well understood and accepted by local patients with heart disease.[10] Protocol compliance was assessed periodically by the study team using quality control criteria developed by the EuroQol Group.[16]

Detailed descriptions of the cTTO and the valuation protocol can be found elsewhere.[17,18] Briefly, the objective of the task was to identify the point of preferential indifference between 10 years of life in the described target state followed by death and a shorter life (x £ 10 years) in full health followed by death. With a defined utility value of 1 for full health, the utility value of the target state can be calculated as $x/10$. For states considered to be worse than death, a lead time of 10 years was added to both alternatives to elicit a negative utility value for the state. The utility value of a worse-than-death health state was calculated as $(x-10)/10$ such that the utility value of each health state is bounded by -1 and 1; 0 represents the value for the "dead" state.

Outcome Measures

EQ-5D-5L

EQ-5D-5L is a generic, multiattribute utility-based instrument. It contains 5 dimensions (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression) and an EQ-VAS of the overall health status.[19] It describes each dimension at 5 levels of severity (broadly corresponding to no problem, slight problems, moderate problems, severe problems, and extreme problems). Thus, it can describe 3125 possible health states. This study used validated Singapore English and Chinese language versions of the EQ-5D-5L.[20,21] EQ-5D-5L has been psychometrically validated for a large number of diseases, including heart disease.[14]

According to the valuation protocol,[17] each participant valued a randomly selected set (called a block) of 10 hypothetical EQ-5D-5L health states. Each block included one very mild health state chosen from 5 prespecified health states (21111, 12111, 11211, 11121, 11112), the most severe health state (55555), and 8 health states chosen from 80 prespecified health states among the remaining 3119 possible health states. Here, the health state "21111" indicated slight problems (level 2 severity) in the first dimension (mobility) and no problems (level 1 severity) in the remaining 4 dimensions. Other health states were defined similarly. The protocol uses 10 blocks of EQ-5D-5L health states, covering a total of 86 unique EQ-5D-5L health states.

It should be noted that the valuation protocol[17] for EQ-5D-5L requires to collect health state preferences using the discrete choice experiment. Nevertheless, preferences using this method were not collected in this study.

HeartQoL

HeartQoL is a heart disease-specific HRQoL instrument.[22] It comprises 14 items with 4 response levels that range from "not bothered" to "bothered a lot." It provides a global score based on the mean values of the responses. The score ranges from 0 (worst HRQoL) to 3 (best HRQoL). HeartQoL has been validated in more than 22 countries. Our study used its official English and Chinese translated versions.

NYHA and CCS functional classifications

The NYHA and CCS classifications are widely used clinical tools that measure cardiac functional capacity and the severity of exertional angina, respectively.[23,24] They classify patients into classes I, II, III, and IV based on limitations because of symptoms (shortness of breath or angina) at various levels of physical activity. A higher class indicates a worse functional capacity. In this study, participants self-evaluated their NYHA and CCS classes based on structured definitions of these classification systems.

Statistical Methods

Sample size

The sample size required to achieve the desired precision of fixed-effect coefficients of health state descriptors in a statistical model estimating utility values using a 20-parameter linear random-effects model was determined. Determination of the sample size was performed using the methodologies proposed by Gandhi et al[25] for the EQ-5D-5L value set studies. A sample size of 400 participants was required to estimate the coefficients with a precision level (95% confidence interval) of ±0.05, considering 0.05 as the minimum important difference for EQ-5D-5L utility values. The other parameters required for the sample size calculation—a residual variance of 0.4 and a design effect of 0.5—were estimated from the EQ-5D-5L value set study in the general Singaporean population.[25] We anticipated that data from 20% of the participants might not be usable (eg, dropouts) and accordingly planned a sample size of 500 participants.

Model development

Various model specifications were explored (see Appendix Tables 1-5 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.010), and the utility values of the resulting models were examined; only the most appropriate models are reported here. In all the models, we defined the dependent variable as disutility (ie, 1 – utility value) for a given health state. Two core models, a 20-parameter linear random-effect model, and an 8-parameter cross-attribute level effects (CALE) model (a constrained nonlinear model), both with random intercepts at the level of individual study participants, and their variants were extensively tested for performance. Because each participant valued 10 health states, participant-specific random-effect intercepts were considered in all the models to account for intraparticipant correlation.

The linear model can be presented as follows:

$$
\begin{aligned}
\text{Linear model(Model 1): } y &= \alpha + \sum_l \sum_d \beta_{dl} X_{dl} + v + e \\
&= \alpha + \beta_{MO2}X_{MO2} + \beta_{MO3}X_{MO3} + \beta_{MO4}X_{MO4} \\
&+ \beta_{MO5}X_{MO5} + \beta_{SC2}X_{SC2} + \beta_{SC3}X_{SC3} + \beta_{SC4}X_{SC4} + \beta_{SC5}X_{SC5} \\
&+ \beta_{UA2}X_{UA2} + \beta_{UA3}X_{UA3} + \beta_{UA4}X_{UA4} + \beta_{UA5}X_{UA5} + \beta_{PD2}X_{PD2} \\
&+ \beta_{PD3}X_{PD3} + \beta_{PD4}X_{PD4} + \beta_{PD5}X_{PD5} + \beta_{AD2}X_{AD2} \\
&+ \beta_{AD3}X_{AD3} + \beta_{AD4}X_{AD5} + v + e
\end{aligned}
$$

where y represents disutility; α, intercept; X_{dl}, fixed-effect indicator variable for the presence of problems on dimension d at level l; β_{dl}, coefficient for the estimated disutility of having problems on dimension d at level l; v, participant-specific random-effect intercept; and e, a residual error.

Given that a preliminary analysis showed nonmonotonicity in coefficients of usual activities for level 4 and 5 in model 1 and in coefficients of pain/discomfort for level 2 and 3 in some other variant of model 1 (explained later) with a significant overlapping of their 95% confidence intervals, each of the β_{dl} coefficients was constrained to have a value greater than or equal to its previous level coefficient β_{dl-1}.

An alternative to the linear model is a nonlinear CALE model. It includes a single coefficient per dimension (b_{MO}, b_{SC}, b_{UA}, b_{PD}, and b_{AD}) representing the disutility of having problems at level 5 and one coefficient for each of levels 2, 3, and 4 (L_2, L_3, L_4), all of which are multiplied by the respective dimensional coefficients. Here, L_l (l = 2, 3, 4) should be interpreted as the ratio of disutility at level l to that at level 5 with disutility at level 5 set to 1. The model assumes that these ratios are constant across all dimensions. Previous studies indicate that the constraint imposed by the multiplicative CALE model is less susceptible to overfitting, reduces the risk of nonmonotonicity, and may have better out-of-sample predictive accuracy than the linear model.[26]

The model can be presented as follows:

$$
\begin{aligned}
\text{CALE model (Model 2): } y &= \alpha + \sum_l \left(\left(\sum_d \beta_d X_{dl} \right) L_l + v + e \right. \\
&= \alpha + (\beta_{MO}X_{MO2} + \beta_{SC}X_{SC2} + \beta_{UA}X_{UA2} + \beta_{PD}X_{PD2} \\
&+ \beta_{AD}X_{AD2})L_2 + (\beta_{MO}X_{MO3} + \beta_{SC}X_{SC3} + \beta_{UA}X_{UA3} \\
&+ \beta_{PD}X_{PD3} + \beta_{AD}X_{AD3})L_3 + (\beta_{MO}X_{MO4} + \beta_{SC}X_{SC4} \\
&+ \beta_{UA}X_{UA4} + \beta_{PD}X_{PD4} + \beta_{AD}X_{AD4})L_4 + (\beta_{MO}X_{MO5} \\
&+ \beta_{SC}X_{SC5} + \beta_{UA}X_{UA5} + \beta_{PD}X_{PD5} + \beta_{AD}X_{AD5}) + v + e
\end{aligned}
$$

where a, X_{dl}, n, *and* e are the same as defined for the linear model.

Each core model was tested with a few variants and their possible combinations. First, core models with an additional term N45 as a fixed effect. The N45 term was defined as an indicator variable for health states having at least one dimension at either level 4 or 5. It is similar to the N3 term used with the EQ-5D-3L value set in the United Kingdom, which may provide additional explanatory value.[27] Second, core models with left-censored the utility values at −1. This variant was considered because participants could hypothetically value a health state lower than −1. Right censoring at 1 was not considered, because 1 is the theoretical upper bound for the utility value of full health. Third, core models with heteroscedastic error term. This variant was considered because the observed variance of the utility values increased with increasing severity of the health states. Heteroscedasticity of the error term was modeled using the log link of a linear regression model with an intercept and 20 indicator variables X_{dl}.

Standard errors, 95% confidence intervals (2.5 and 97.5 percentiles), and 1-sided *P* values for all the model coefficients were calculated using bootstrap sampling (1000 participant-level samples).

Model selection

The predictive accuracy of the models was evaluated in terms of mean absolute error (MAE), root mean square error (RMSE), and Lin's concordance coefficient between the predicted and mean values of the observed values of the health state. Lower MAE and RMSE and higher concordance coefficient indicate better predictive accuracy. The out-of-sample fit was evaluated in cross-validation samples. Cross-validation was performed by fitting the models to a subset of the data set prepared by excluding one of the 10 blocks of health states and assessing the predictive accuracy in the excluded block.[26] In-sample fit was assessed using the full data set. For selecting the preferred model, priority was given to out-of-sample fit predictive accuracy as measured using RMSE, followed by in-sample fit predictive accuracy, the model using the least number of fixed-effect parameters, and achieving the lowest Bayesian information criterion (ie, model parsimony).

Rescaling

The predicted utility value for full health ("11111") may not be 1 because of the nonzero intercept in the preferred model. We rescaled all the predicted utility values by dividing them with 1 – intercept to obtain a value of 1 for full health and proportionally adjusted values of the other health states.[28]

All models were fitted using the xreg package[29] for R software.[30]

Model validation

The preferred model was assessed for the known-groups discriminative ability of its predicted utility values (rescaled). Mean utility values based on the participants' own EQ-5D-5L health states were estimated for each of the NYHA and CCS functional classes and for the EQ-VAS and HeartQoL global score classes. EQ-VAS and HeartQoL global scores were divided into 3 classes using their first (Q1) and third (Q3) quartiles (class I, ≥Q3; class II, Q1-Q3; class III, £Q1). Lower class represents better health or functional capacity. Mean utility values were expected to decrease as class increased. Mean utility values across the classes and differences between 2 individual classes were compared using analysis of variance and the two-sample *t* test, respectively.

Results

A total of 1166 potential participants were approached for this study. Of these, 64% were willing to participate, and 78% of those met the eligibility criteria. Of the recruited participants (N = 582), 6 were excluded from analysis: 3 were recruited twice during the quality control review, and 3 did not complete the interview (Appendix Fig. 1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.010). Therefore, 576 participants were included in the analysis. The sociodemographic and health characteristics of the participants who were included in the analysis are presented in Table 1. The study recruited older, more men, and Malay and Native American participants than those found in the Singapore population.[15]

Among 5760 cTTO responses, 2359 responses (41.0%) were considered worse than death (see Appendix Fig. 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.010). The proportion of values clustered at 1, 0, and −1 was 5.4%, 6.2%, and 18.0%, respectively. There was negative correlation between the mean utility values and misery score (sum of the severity levels across all 5 dimensions) for both worse-than-death values (correlation −0.81; *P*<.001) and better-than-death values (correlation −0.95, *P*<.001) (Appendix Fig. 3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.010).

In general, there was no or trivial improvement in predictive accuracy with heteroscedastic residual error models compared with the similar models without heteroscedastic error. In addition, the models with left censoring had systematically much lower predicated values than the observed mean values for both better-than-death and worse-than-death health states (see Appendix Tables 1-5 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.010). Hence, the models without heteroscedastic residual error and without left censoring were considered for further scrutiny.

Among the other linear models (without left censoring and without heteroscedastic residual error), the model with the N45 term had lower MAE and RMSE and higher concordance coefficients with both the full and cross-validation data sets than the model without the N45 term. The coefficient of the N45 term was also statistically significant (*P*<.001) in the model (see Appendix Tables 1-3 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.010). As in the linear models, inclusion of the N45 term improved the performance of the CALE model (without left censoring and without heteroscedastic residual error), and the coefficient for the N45 term was statistically significant (*P*<.001) (Appendix Tables 1, 4, and 5 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.010). Although the predictive accuracy (MAE, RMSE, concordance coefficient) of the CALE model was comparable with the linear model with the N45 term for the out-of-sample fit, it was lower than that of the linear model for in-sample fit (Appendix Fig. 4 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.010). Hence, the linear model with the N45 term (model 1 + N45 term) without heteroscedastic error and without left censoring was selected as the preferred model for developing the value set. Table 2 compares the performances of linear and CALE models (without heteroscedastic residual error and without left censoring) using both the full and cross-validation data sets. Figure 1 shows the predicted utility values obtained using the linear and CALE models with the N45 term (without heteroscedastic residual error and without left censoring) as a function of the directly valued health states of the participants.

Coefficients for the preferred model are presented in Table 3. The largest and smallest mean utility decrements were for

Table 1. Patient characteristics.

Characteristics	Our study (N = 576)	General population,* %
Age (years), mean (SD)	57.7 (11.5)	
21-40, n (%)	45 (7.8)	40.9
41-60, n (%)	273 (47.4)	42.5
>60, n (%)	258 (44.8)	16.6
Men, n (%)	416 (72.2)	49.1
Ethnicity, n (%)		
Chinese	328 (56.9)	75.8
Malay	119 (20.7)	12.1
Native American	97 (16.8)	8.8
Others	32 (5.6)	3.2
Educational level, n (%)		
Primary (6 years) or les	101 (17.5)	8.2
Secondary (up to 11 years)	271 (47.1)	34.2
Diploma, university, or higher	204 (35.4)	56.6
Married, n (%)	423 (73.4)	65.8
Monthly household income of <S$4000, n (%)	292 (50.7)	37.5
Employed, n (%)	324 (56.3)	
Heart disease diagnosis,† n (%)		
Ischemic heart disease	456 (79.2)	
Heart rhythm disorder	167 (29.0)	
Heart failure	157 (27.3)	
Valvular heart disease	97 (16.8)	
Other heart problems	33 (5.7)	
Number of comorbidities, n (%)		
0	62 (10.8)	
1-2	228 (39.6)	
3-4	238 (41.3)	
>4	48 (8.3)	
NYHA functional classification		
I	268 (46.5)	
II	254 (44.1)	
III-IV	54 (9.4)	
CCS functional classification for angina		
I	453 (78.7)	
II	104 (18.1)	
III-IV	19 (3.3)	
EQ-VAS, mean (SD)	77.2 (14.9)	
HeartQoL global score, mean (SD)	2.35 (0.55)	

CCS indicates Canadian Cardiovascular Society; EQ-VAS, EQ visual analog scale; NYHA, New York Heart Association.
*General population at the age of 20 to 79 years based on Singapore census 2010.[15]
†A patient may have multiple heart disease diagnoses; hence, he/she may be counted under more than 1 diagnosis.

mobility and anxiety/depression dimensions, respectively. This was also the case for the CALE model with the N45 term, indicating that mobility and anxiety/depression dimensions have the highest and lowest impacts, respectively, on disutility values (see Appendix Table 5 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.010).

Utility values predicted by the preferred model were rescaled by dividing each value by 1 – intercept = 1 – 0.196 = 0.804 for the value set. The value set has values of 1, 0.981, and −0.727 for full health, second-best ("11112"), and the worst state ("55555"), respectively. Figure 2 shows the originally predicted and rescaled utility values for the preferred model. An example that demonstrates how to use the coefficients of the preferred model to calculate the utility values with and without rescaling can be found in Appendix Table 6 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.010. The utility values with

Table 2. Comparison of model performance.

Predictive accuracy measures	Linear models		CALE models	
	Model 1	Model 1 + N45 (preferred model)	Model 2	Model 2 + N45
Mean absolute error				
Full data set	0.075	0.063	0.082	0.070
Cross-validation data set	0.095	0.081	0.092	0.081
Root mean square error				
Full data set	0.093	0.079	0.100	0.089
Cross-validation data set	0.118	0.101	0.116	0.105
Concordance coefficient				
Full data set	0.972	0.980	0.968	0.975
Cross-validation data set	0.956	0.968	0.957	0.965
Number of fixed-effect parameters	21	22	9	10
BIC based on the full data set	7745.8	7657.9	7680.8	7605.1

Note. Model 1, 20-parameter linear random-effect model. Model 2, 8-parameter CALE model. N45, indicator variable for states with at least one dimension at a severity level of either 4 or 5. See Methods section for details.
BIC indicates Bayesian information criterion; CALE, cross-attribute level effects.

and without rescaling for all 3125 health states are available in Appendix 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.09.010.

The mean utility values based on the preferred model using the participants' own health states ranked in the expected high to low direction for participants with increasing NYHA and CCS classes I to III/IV ($P<.001$) (Table 4). Differences in mean utility value between 2 consecutive classes were also statistically significant ($P<.01$), with most mean differences of $\geq$0.05 (minimum important difference). Similar results were observed for EQ-VAS and HeartQoL classes (Table 4).

Discussion

A value set for EQ-5D-5L using preferences of patients with heart disease was developed according to a standardized international protocol. This is the first EQ-5D-5L value set developed using patient preferences exclusively. For patients with heart

Figure 1. Observed and predicted utility values for directly valued health states. Utility values are sorted based on observed mean values.

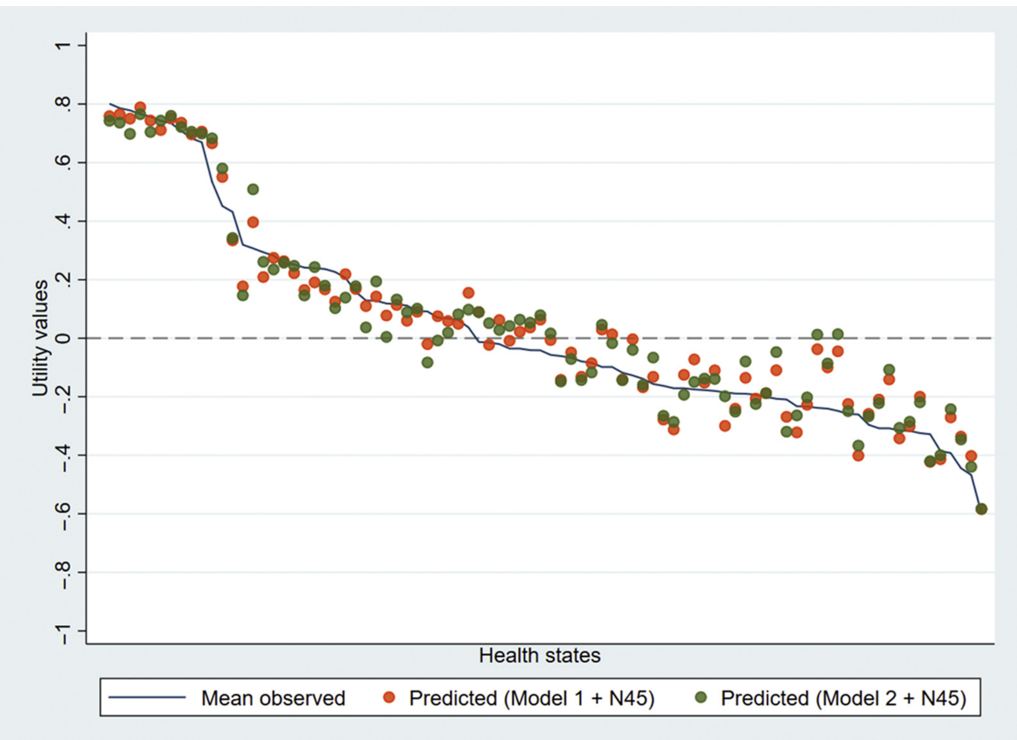


Table 3. Summary of the preferred model (model 1 + N45) based on the full data set.

Variables	Coefficient	Standard error	95% Confidence interval	*P* value
Intercept	0.196	0.030	0.142-0.253	<.001
MO2	0.039	0.020	0.000-0.078	.024
MO3	0.106	0.020	0.065-0.145	<.001
MO4	0.200	0.023	0.155-0.244	<.001
MO5	0.281	0.020	0.241-0.319	<.001
SC2	0.092	0.018	0.057-0.127	<.001
SC3	0.193	0.020	0.156-0.230	<.001
SC4	0.246	0.021	0.203-0.286	<.001
SC5	0.273	0.018	0.237-0.308	<.001
UA2	0.052	0.019	0.014-0.088	.002
UA3	0.121	0.022	0.074-0.164	<.001
UA4	0.169	0.018	0.134-0.204	<.001
UA5	0.172	0.017	0.138-0.207	<.001
PD2	0.045	0.016	0.014-0.075	.002
PD3	0.055	0.018	0.022-0.091	.001
PD4	0.228	0.018	0.193-0.262	<.001
PD5	0.239	0.018	0.204-0.278	<.001
AD2	0.015	0.015	0.000-0.048	.155
AD3	0.094	0.022	0.053-0.135	<.001
AD4	0.116	0.021	0.076-0.157	<.001
AD5	0.167	0.019	0.133-0.208	<.001
N45	0.255	0.028	0.200-0.311	<.001

Note. Model 1, 20-parameter linear random-effect model (see Methods section for details). MO2 to MO5, SC2 to SC5, UA2 to UA5, PD2 to PD5, and AD2 to AD5 represent indicator variables for severity levels 2 to 5 with reference to level 1 for mobility, self-care, usual activities, pain/discomfort, and anxiety/depression dimensions, respectively. N45 represents an indicator variable for health states with at least one dimension at level 4 or 5.

disease, it can inform patient-centric economic evaluations and clinical decision making and be used to evaluate differences in the outcomes of such decisions derived using societal versus patient preferences. The utility index based on the preferences of patients with heart disease also demonstrated known-group validity in differentiating patients with different levels of disease severity. The utility index can also be used as a HRQoL measure in clinical research of patients with heart disease.

We have presented utility values with and without rescaling for the developed value set. Rescaling was used because of a nontrivial intercept value (ie, 0.196). A nontrivial intercept results in a sizable difference between the full health value (which is set

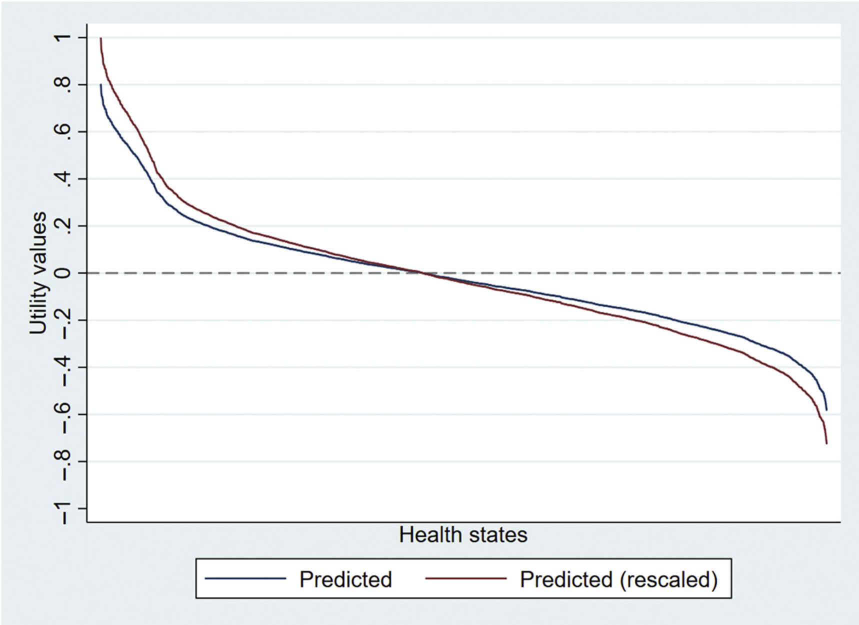
Table 4. Known-group validity of rescaled utility values derived from the preferred model (model 1 + N45).

Known groups	NYHA		CCS		EQ-VAS		HeartQoL global	
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
Class I	268	0.960 (0.093)	453	0.949 (0.106)	161	0.973 (0.048)	164	0.987 (0.031)
Class II	254	0.926 (0.123)	104	0.840 (0.251)	202	0.932 (0.153)	200	0.955 (0.069)
Class III/IV	54	0.672 (0.372)	19	0.596 (0.377)	213	0.863 (0.227)	212	0.829 (0.252)
ANOVA *P* value		<.001		<.001		<.001		<.001

	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)
Diff (I-II)	0.034* (0.015-0.052)	0.110* (0.079-0.140)	0.042† (0.017-0.066)	0.031* (0.020-0.043)
Diff (II-III/IV)	0.254* (0.198-0.310)	0.244* (0.109-0.379)	0.069* (0.032-0.107)	0.126* (0.090-0.163)
Diff (I-III/IV)	0.288* (0.237-0.339)	0.353* (0.294-0.412)	0.111* (0.075-0.147)	0.158* (0.119-0.197)

Note. Classes I, II and III/IV for EQ-VAS and HeartQoL global represent ≥ Q3, Q1 to Q3, and ≤ Q1 of their values, respectively.
ANOVA indicates analysis of variance; CCS, Canadian Cardiovascular Society; CI, confidence interval; Diff, difference; EQ-VAS, EQ visual analog scale; NYHA, New York Heart Association; Q1, first quartile; Q3, third quartile.
*2-sample *t* test $P<.001$.
†2-sample *t* test $P<.01$.

Figure 2. Predicted utility values using the preferred model (model 1 + N45) for all possible health states. Utility values are sorted based on predicted values.



at 1) and the next best health state value (ie, 0.789). Such values mean that a change from full health to 11112 is associated with a drop in utility as big as >20% of the 0 (dead) to 1 (full health) scale, suggesting poor face validity. Rescaling resolves this issue by adjusting the values for both full health and all other health states to preserve their relative distance as predicted by the model. Rescaled values would minimize chances of promoting healthcare interventions that focus on mostly improving mild health problems and favor lifesaving interventions. Therefore, we recommend the use of the rescaled values.

Mobility is the most relevant EQ-5D dimension, and anxiety/depression the least, in terms of impact on utility value decrement according to preferences of patients with heart disease. This differs from the preferences of the Singapore general population, which considers usual activities the most relevant dimension of the EQ-5D-3L value set and pain/discomfort the least.[31] Such differences are expected. Mobility is essential to life in Singapore, where most people work past official retirement and commute on public transport. Heart disease often significantly limits physical activity such as walking, which explains patients' preference for avoiding this dimension. That anxiety/depression is the least important dimension might be related to mental adaptation to disease, because most heart diseases are chronic. Previous studies have also found differences between the preferences of patients with heart disease and the general population.[10-12] Our results provide granularity regarding the dimensions in which the preferences differ.

We chose a 20-parameter linear model with the N45 term as the final model for developing a value set. We observed logical inconsistencies in the initial version of the model and had to constrain coefficients to achieve monotonicity. Such logical inconsistencies have also been observed in several countries' value sets for the EQ-5D-5L.[26,32-34] This could be due to the complexity of the model, which might predispose patients to overfitting to random variance. Nevertheless, the constrained linear model still provided better predictive accuracy on the full-sample than the constrained 8-parameter CALE models, possibly because the assumption of a constant ratio of level parameters across the dimensions was not fully satisfied in our study data.

There were demographic differences between the study and general populations. The former included more elderly individuals, men, and individuals with lower educational levels and had a higher representation of Malay and Native American ethnicities than the latter. These characteristics are known risk factors for heart disease in Singapore.[35] Hence, a higher representation of these characteristics in the patient sample is not unexpected and supports the sample's face validity.

There have been attempts to develop disease-specific value sets for health dimensions affected by specific diseases or their treatments. For example, a cancer-specific QLU-C10D descriptive system has been developed from the EORTC C30 HRQoL measure,[36] and its value sets have been or are being developed in several countries.[37,38] We chose the EQ-5D-5L for our study for 2 reasons: (1) to the best of our knowledge, no heart disease-specific descriptive system that can be used to develop a preference-based value set is currently available, and (2) the EQ-5D-5L and its former version the EQ-5D-3L have demonstrated acceptable measurement properties in patients with heart disease and are widely used for economic evaluation using societal preferences. A value set based on patient preferences for the same descriptive system can facilitate comparisons of economic evaluations based on patients' and societal perspectives. Unfortunately, there is no Singaporean general public value set for EQ-5D-5L yet. Nevertheless, a previous study[10] that used a limited set of EQ-5D-5L health states in Singapore found that utility gain could be higher for interventions improving health status from severe to mild or moderate using the heart disease patient values than the general public values, but such gain might not be apparent for interventions that could improve the health status from moderate to mild if the heart disease patient values are used. Therefore, we expect that differences in the outcome of economic evaluation also depend on health states observed in the study. Further research to compare the present value set for patients with heart disease and the Singaporean general public value set will be needed when the latter becomes available.

Our study has some limitations. Given that it would have been difficult to conduct cognitively demanding valuation tasks among hospitalized patients, we could only approach patients in outpatient clinics. Nevertheless, we enriched the sample by recruiting those with prior hospitalizations. Recruiting participants with heart disease among the general population would be ideal but is logistically challenging. We believed that recruiting participants from hospital outpatient clinics would help us sample the target population (patients with documented clinical diagnoses based on hospital records) without screening a large number of "generally healthy" candidates. Notably, the profiles of the study patients are comparable with patients with heart disease recruited from the Singapore general population in Gandhi et al,[11] which suggests our study findings can be generalizable. The requirement for basic literacy to complete the valuation tasks could have selectively excluded some participants, especially the elderly, who would otherwise have qualified but is common in most valuation studies. We reported preferences of patients in Singapore but patient preferences can vary among countries, possibly because of differences in culture and healthcare systems. Therefore, the appropriateness of this value set should be evaluated before adoption in other countries. The current study has not explored potential impact of different types of heart diseases, comorbidities, length of the disease, treatment, and socioeconomic characteristics on utility values of different severity. This should be studied in the future to understand whether patient characteristics affect their preferences and, if so, to what extent.

Conclusions

We have developed a time trade-off–based EQ-5D-5L value set using the preferences of patients with heart disease that enables patient-centric HTAs and clinical decision making for treatment selection. The value set provides a new EQ-5D-5L index for measuring the HRQoL of patients with heart disease.

Supplemental Materials

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2021.09.010.

Article and Author Information

Accepted for Publication: September 6, 2021

Published Online: November 25, 2021

doi: https://doi.org/10.1016/j.jval.2021.09.010

Author Affiliations: Biostatistics, Singapore Clinical Research Institute, Singapore (Gandhi); Centre for Quantitative Medicine, Duke-NUS Medical School, Singapore (Gandhi, Cheung); Centre for Child Health Research, Tampere University, Finland (Gandhi, Cheung); Department of Cardiology, National Heart Centre Singapore, Singapore (Tan, Lam); Department of Cardiology, National University Heart Centre, Singapore (Lim); Health Services Research Centre, Akershus University Hospital, Lørenskog, Norway (Rand); Maths in Health B. V, Rotterdam, The Netherlands (Rand); Saw Swee Hock School of Public Health, National University of Singapore, Singapore (Luo).

Correspondence: Mihir Gandhi, PhD, Biostatistics, Singapore Clinical Research Institute, #06-01, 23 Rochester Park, Singapore 139234. Email: mihir.gandhi@scri.cris.sg

Author Contributions: *Concept and design:* Gandhi, Tan, Rand, Lam, Luo, Cheung
Acquisition of data: Gandhi, Tan, Lim
Analysis and interpretation of data: Gandhi, Tan, Lim, Rand, Lam, Luo, Cheung
Drafting of the manuscript: Gandhi
Critical revision of the paper for important intellectual content: Gandhi, Tan, Lim, Rand, Lam, Luo, Cheung
Statistical analysis: Gandhi, Rand, Luo, Cheung
Provision of study materials or patients: Tan, Lam, Lim, Luo
Obtaining funding: Rand, Luo, Cheung
Administrative, technical, or logistic support: Gandhi, Tan, Lim
Supervision: Luo, Cheung

Conflict of Interest Disclosures: Drs Gandhi, Rand, and Luo reported receiving support for attending meetings and/or travel from the EuroQol Research Foundation. Drs Rand and Luo reported receiving a grant from the EuroQol Research Foundation for this study that was paid to the principal investigator's institution. Dr Rand is currently the chairman of the EuroQol Executive Committee and has received funding from the EuroQol Research Foundation for various other research projects related to HRQoL instruments. Dr Lam reported receiving grants from AstraZeneca, Bayer, Boston Scientific, and Roche Diagnostics; consulting fees from Actelion, Amgen, Applied Therapeutics, AstraZeneca, Bayer, Boehringer Ingelheim, Boston Scientific, Cytokinetics, Darma Inc, Janssen R&D, Medscape/WebMD Global, Merck, Novartis, Novo Nordisk, Roche Diagnostics, Sanofi, St Luke, and Us2.ai; payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing, or educational events from Amgen, AstraZeneca, Bayer, Boehringer Ingelheim, Medscape/WebMD Global, Novartis, Radcliffe, Roche Diagnostics; and a patent pending (PCT and SG 2016/050217) and a patent application (16/216,929); participated on a data safety monitoring board or advisory board for Amgen, Astra Zeneca, Bayer, Boston Scientific, Novartis, Novo Nordisk, and Roche Diagnostics; reported holding a leadership role (nonexecutive director); and reported holding stocks and stock options in Us2.ai and being a cofounder of Us2.ai. Dr Cheung reported receiving grants paid to his institution from the Agency for Science, Technology and Research (A*STAR), Singapore, and the Ministry of Health, Singapore. Dr Luo is an editor for *Value in Health* and had no role in the peer-review process of this article. No other disclosures were reported.

Funding/Support: The study was jointly funded by the Duke-National University of Singapore Medical School Signature Research Program (funded by the Agency for Science, Technology and Research (A*STAR), Singapore, and the Ministry of Health, Singapore; account R-913-200-040-263) and the EuroQol Research Foundation (EQ Project 20170430).

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Acknowledgment: The authors thank Tan Gretchen, Woo Kai Lee, and Koh Rui Xuan Audrey for their contribution to the data collection. The Clinical Trials Research Office, National Heart Centre Singapore, has provided administrative support.

REFERENCES

1. World population ageing 2017. Union Nations, Department of Economic and Social Affairs Population Division. https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017_Report.pdf. Accessed January 2021.
2. Drummond M, Torbica A, Tarricone R. Should health technology assessment be more patient centric? If so, how? *Eur J Health Econ*. 2020;21(8):1117–1120.
3. 2015 global survey on health technology assessment by national authorities. World health organization. https://www.who.int/health-technology-assessment/MD_HTA_oct2015_final_web2.pdf. Accessed January 15, 2021.
4. Helgesson G, Ernstsson O, Åström M, Burström K. Whom should we ask? A systematic literature review of the arguments regarding the most accurate source of information for valuation of health states. *Qual Life Res*. 2020;29(6):1465–1482.
5. Cubi-Molla P, Shah K, Burström K. Experience-based values: a framework for classifying different types of experience in health valuation research. *Patient Patient Centered Outcomes Res*. 2018;11(3):253–270.
6. General guidelines for economic evaluations from the Dental and Pharmaceutical Benefits Agency (TLVAR 2017:1). TLV. https://www.tlv.se/

download/18.467926b615d084471ac3230c/1510316374332/TLVAR_2017_1.pdf. Accessed January 2021.
7. Drug Evaluation Methods and Process Guide. 2nd version ACE. https://www.ace-hta.gov.sg/docs/default-source/process-methods/ace-methods-and-process-guide-for-drug-evaluation-(20-dec-2019).pdf. Accessed January 2021.
8. Medical Technologies Methods and Process Guide. 1st version. ACE. http://www.ace-hta.gov.sg/public-data/our-process-and-methods/ACE%20methods%20and%20process%20guide%20for%20medical%20technologies%20evaluation%20(1%20Oct%202018).pdf. Accessed January 2021.
9. World health statistics. World Health Organization. http://apps.who.int/iris/bitstream/10665/112738/1/9789240692671_eng.pdf. Accessed January 2021.
10. Gandhi M, San Tan R, Ng R, et al. Comparison of health state values derived from patients and individuals from the general population. *Qual Life Res*. 2017;26(12):3353–3363.
11. Gandhi M, Thumboo J, Luo N, Wee HL, Cheung YB. Do chronic disease patients value generic health states differently from individuals with no chronic disease? A case of a multicultural Asian population. *Health Qual Life Outcomes*. 2015;13(1):8.
12. Pickard AS, Tawk R, Shaw JW. The effect of chronic conditions on stated preferences for health. *Eur J Health Econ*. 2013;14(4):697–702.
13. Roth GA, Johnson C, Abajobir A, et al. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990-2015. *J Am Coll Cardiol*. 2017;70(1):1–25.
14. Janssen M, Pickard AS, Golicki D, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res*. 2013;22(7):1717–1727.
15. Census of population 2010 statistical release 1. Demographic characteristics, education, language and religion. Singapore Department of Statistics. https://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications_and_papers/cop2010/census_2010_release1/cop2010sr1.pdf. Accessed January 2021; Published 2010.
16. Ramos-Goñi JM, Oppe M, Slaap B, Busschbach JJ, Stolk E. Quality control process for EQ-5D-5L valuation studies. *Value Health*. 2017;20(3):466–473.
17. Oppe M, Devlin NJ, van Hout B, Krabbe PF, de Charro F. A program of methodological research to arrive at the new international EQ-5D-5L valuation protocol. *Value Health*. 2014;17(4):445–453.
18. Oppe M, Rand-Hendriksen K, Shah K, Ramos-Goñi JM, Luo N. EuroQol protocols for time trade-off valuation of health outcomes. *Pharmacoeconomics*. 2016;34(10):993–1004.
19. EQ-5D-5L user guide. EuroQol Research Foundation. https://euroqol.org/wp-content/uploads/2021/01/EQ-5D-5LUserguide-08-0421.pdf. Accessed May 2021.
20. Luo N, Wang Y, How CH, Tay EG, Thumboo J, Herdman M. Interpretation and use of the 5-level EQ-5D response labels varied with survey language among Asians in Singapore. *J Clin Epidemiol*. 2015;68(10):1195–1204.
21. Luo N, Wang Y, How C, et al. Cross-cultural measurement equivalence of the EQ-5D-5L items for English-speaking Asians in Singapore. *Qual Life Res*. 2015;24(6):1565–1574.
22. Oldridge N, Höfer S, McGee H, et al. The HeartQoL: Part I. Development of a new core health-related quality of life questionnaire for patients with ischemic heart disease. *Eur J Prev Cardiol*. 2014;21(1):90–97.
23. Dolgin M, Fox A, Gorlin R, Levin R, New York Heart Association, Criteria Committee. *Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels*. 9th ed. Boston, MA: Lippincott Williams and Wilkins; 1994.
24. Campeau L. Grading of angina pectoris. *Circulation*. 1976;54(3):522–523.
25. Gandhi M, Xu Y, Luo N, Cheung YB. Sample size determination for EQ-5D-5L value set studies. *Qual Life Res*. 2017;26(12):3365–3376.
26. Rand-Hendriksen K, Ramos-Goñi JM, Augestad LA, Luo N. Less is more: cross-validation testing of simplified nonlinear regression model specifications for EQ-5D-5L health state values. *Value Health*. 2017;20(7):945–952.
27. Dolan P. Modeling valuations for EuroQol health states. *Med Care*. 1997;35(11):1095–1108.
28. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. *Value Health*. 2017;20(4):662–669.
29. Rand K. xreg: flexible, multi-data.frame maximum likelihood-based regression functions in R. GitHub. https://github.com/intelligentaccident/xreg. Accessed January 2021.
30. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org. Accessed January 2021.
31. Luo N, Wang P, Thumboo J, Lim Y-W, Vrijhoef HJ. Valuation of EQ-5D-3L health states in Singapore: modeling of time trade-off values for 80 empirically observed health states. *Pharmacoeconomics*. 2014;32(5):495–507.
32. Xie F, Pullenayegum E, Gaebel K, et al. A time trade-off-derived value set of the EQ-5D-5L for Canada. *Med Care*. 2016;54(1):98.
33. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5 D-5 L value set for E ngland. *Health Econ*. 2018;27(1):7–22.
34. Versteegh MM, Vermeulen KM, Evers SM, de Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. *Value Health*. 2016;19(4):343–352.
35. Picco L, Subramaniam M, Abdin E, Vaingankar JA, Chong SA. Prevalence and correlates of heart disease among adults in Singapore. *Asian J Psychiatry*. 2016;19:37–43.
36. King MT, Costa D, Aaronson N, 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):625–636.
37. King MT, Viney R, Pickard AS, et al. Australian utility weights for the EORTC QLU-C10D, a multi-attribute utility instrument derived from the cancer-specific quality of life questionnaire, EORTC QLQ-C30. *Pharmacoeconomics*. 2018;36(2):225–238.
38. Norman R, Mercieca-Bebber R, Rowen D, et al. UK utility weights for the EORTC QLU-C10D. *Health Econ*. 2019;28(12):1385–1401.