



Estimation of an EORTC QLU-C10 Value Set for Spain Using a Discrete Choice Experiment

Aureliano Paolo Finch^{1,7} · Eva Gamper² · Richard Norman³ · Rosalie Viney⁴ · Bernhard Holzner^{2,5} · Madeleine King⁶ · Georg Kemmler⁵ on behalf of the EORTC Quality of Life Group

Accepted: 7 June 2021 / Published online: 3 July 2021
© The Author(s) 2021

Abstract

Background The EORTC QLU-C10D is a preference-based measure derived from the EORTC QLQ-C30. For use in economic evaluations, country-specific value sets are needed. This study aimed to generate an EORTC QLU-C10 value set for Spain.

Methods A sample of the Spanish general population completed an online discrete choice experiment. An attribute-balanced incomplete block design was used to select 960 choice tasks, with a total of 1920 health states. Each participant was randomly assigned 16 choice sets without replacement. Data were modelled using generalized estimating equations and mixed logistic regressions.

Results A total of 1625 panel members were invited to participate, 1010 of whom were included in the study. Dimension decrements were generally monotonic with larger disutilities at increased severity levels. Dimensions associated with larger decrements were physical functioning and pain, while the dimension with the smallest decrement was sleep disturbances. The PITS state (i.e. worst attainable health) for the Spanish population is – 0.043.

Conclusions This study generated the first Spanish value set for the QLU-C10D. This can facilitate cost-utility analyses when applied to data collected with the EORTC QLQ-C30.

✉ Aureliano Paolo Finch
finc8@hotmail.com

- ¹ Research Centre on Health and Social Care Management (CERGAS), Bocconi University, Via Sarfatti 25, S1 4DT Milan, Italy
- ² Division of Psychiatry I, Department of Psychiatry, Psychotherapy and Psychosomatics, Medical University of Innsbruck, Innsbruck, Austria
- ³ School of Public Health, Curtin University, Perth, Australia
- ⁴ Centre for Health Economics Research and Evaluation (CHERE), UTS Business School, University of Technology Sydney (UTS), Sydney, NSW, Australia
- ⁵ Division of Psychiatry II, Department of Psychiatry, Psychotherapy and Psychosomatics, Medical University of Innsbruck, Innsbruck, Austria
- ⁶ School of Psychology, University of Sydney, Sydney, Australia
- ⁷ Health Values Research and Consultancy, Amsterdam, The Netherlands

Key Points for Decision Makers

This study provides the first value set for the EORTC QLU-C10D for Spain. This preference-based measure is derived from the widely used cancer-specific quality-of-life questionnaire, EORTC QLQ-C30.

Cost-utility analysis (CUA) is an important component of reimbursement decisions in many countries, including Spain. The availability of the EORTC QLU-C10D value set supports CUA for cancer interventions in Spain and enables cross-country comparisons.

1 Introduction

Cost-utility analysis (CUA) is often used to inform whether new treatments or interventions should be reimbursed within a healthcare system. CUA typically measures benefits in

terms of quality-adjusted life-years (QALYs), combining the length of life with an index of health-related quality of life (HRQoL), often known as a utility, or values, of patients. Values can be obtained using a variety of direct and indirect methods, including using description of health states associated with a condition (i.e. vignettes), asking patients to value their own health directly or, more commonly, indirect valuation using a preference-based measure (PBM) of health [1].

PBMs comprise a descriptive system through which health is described, and a value set reflecting strength of preference of members of the general public, or more rarely patients, for the health state described. PBMs can be generic (GPBMs) or condition-specific. While GPBMs can be used to describe health problems relevant across multiple diseases and conditions, condition-specific PBMs typically describe health problems occurring in a specific condition of interest [1–3], often focussing on particular symptoms, aspects of functioning or side effects.

For use in CUA, prominent health technology assessment bodies such as the National Institute for Health and Care Excellence (NICE) in the UK and the Zorginstituut in the Netherlands express a preference for using GPBMs [4, 5]. In other countries, for example in Spain, the choice of which measure to use for obtaining values for QALYs calculation is left to the analyst, who can consider either GPBM or condition-specific PBM [6]. There are multiple arguments supporting the use of GPBM, including the fact that GPBMs detect the negative impact of a wide range of comorbidities alongside the positive impact of interventions, as well as avoiding labelling and focusing effects biases [2]. More importantly, using a single GPBM for all assessments across all conditions allows for cross-program comparability [7]. The potential concern with using a single GPBM alone is that the chosen GPBM may not be valid and responsive for the group of patients being examined. In those cases, using a different GPBM or condition-specific PBM is necessary [5].

A limited number of GPBMs dominate the literature [8], specifically the EQ-5D, the SF-6D, the Health Utility Index mark 3 (HUI3), the Assessment of Quality of Life (AQoL), the 15 Dimensions (15D) and the Quality of Well-being Self-Administered (QWB-SA) [9, 10]. Although there is considerable evidence for their validity and responsiveness in many common medical problems such as skin, respiratory, genitourinary, endocrine, nutritional and metabolic diseases [7, 11, 12], there is also mixed or inconsistent evidence for all these measures in some specific types of cancers [7, 11, 13].

Recently, a partnership between the Multi-Attribute Utility in Cancer (MAUCa) Consortium and the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Group led to the development of the EORTC QLU-C10D [14]. The EORTC QLU-C10D is a condition-specific PBM derived from the EORTC

QLQ-C30, one of the most commonly used patient-reported outcome measures in cancer randomized controlled trials [15]. Given the widespread use of the EORTC QLQ-C30, the EORTC QLU-C10D might be an alternative when the preferred GPBM is not included in the trial of interest, or a useful measure to perform CUA sensitivity analysis in those cases in which the preferred GPBM reports mixed validity and responsiveness results. This is because the QLU-C10D allows values to be directly estimated from clinical studies that have used the QLQ-C30, without the need for mapping onto GPBMs or additional data collection.

The development of the EORTC QLU-C10D followed two stages. First, the health state classification system was developed to reduce the thirty items of the EORTC QLQ-C30 into 13 items covering ten dimensions [16]. Subsequently, a valuation method using a discrete choice experiment (DCE) was developed [17], which was then used to generate a value set from members of the general public of Australia [18]. This valuation method is valid and has been increasingly used due to its ease of application and reduction in data collection burden [19, 20].

For the conduct of CUAs, it is recommended to use country-specific value sets, as differences in preferences across countries might substantially alter the values obtained [21–23] and consequently cost-effectiveness estimates. Hence, valuation studies for the EORTC QLU-C10D have been completed, or are currently being undertaken, in numerous countries, including Canada [24], Germany [14], the United Kingdom [25], Austria, France, Italy, Poland [26], the Netherlands, the United States [14] and Singapore [27]. The current study reports on the EORTC QLU-C10D valuation in Spain. Preferences were elicited from a representative sample of the Spanish general population, replicating the methods employed by King and colleagues [18].

2 Methods

2.1 EORTC QLQ-C30 and QLU-C10D

The EORTC QLQ-C30 is a non-preference-based HRQoL measure widely used in cancer [28]. Its 30-item descriptive system covers the most common cancer symptoms, such as fatigue, pain and side effects of chemotherapy (e.g. nausea, vomiting etc.) and important aspects of functioning (e.g. physical, cognitive, emotional, role etc.). The measure is summarized using subscales related to a symptom or functioning aspect, and by a global quality-of-life scale.

The EORTC QLU-C10D was developed by King and colleagues with the objective of ensuring a descriptive system amenable for valuation [16]. It covers 10 key EORTC QLQ-C30 dimensions, namely physical functioning, role functioning, social functioning, emotional functioning, pain,

fatigue, sleep disturbances, appetite loss, nausea and bowel problems. Each item has four possible levels of severity. Hence, the measure can describe over a million possible health states ($4^{10} = 1,048,576$). Table 1 presents the classification system of the QLU-C10D and how it maps to the 13 component items from the QLQ-C30. Of note, the EORTC QLU-C10 is not a brief form profile instrument, nor a standalone measure, but a utility scoring algorithm that can be used in trials employing the EORTC QLQ-C30.

2.2 Discrete Choice Experiment (DCE) Design and Randomization

The QLU-C10D Spanish valuation used the same DCE design tested in Norman and colleagues [17] and employed for the Australian valuation in King et al. [18]. Broadly, the design is based on 960 choice sets selected to ensure statistical efficiency for the model's parameter estimation based on the results of a design experiment [17]; this is described in more detail below.

Each choice set presents two different health states, each of which is described in terms of the 10 health domains of the QLU-C10D (the HRQoL attributes in the DCE) and a duration attribute. In the DCE, the physical functioning domain was presented as two attributes (related to short walk and long walk), to simplify the wording of the attribute [18]. However, the underlying design treats it as a single attribute. To reduce responders' cognitive burden, only four of the QLU-C10D attributes and the duration attribute differed in each choice set, while the remaining were the same ('overlapped') across pairwise choices.

Selection of the health states began with the determination of a balanced incomplete block design where the four attributes that varied across each health state were identified. This balanced incomplete block design was then duplicated for the other dimensions. A generator-based approach was used to determine which levels of the four attributes allowed the estimation of main effects and two-level interactions involving duration [29]. In order to determine the levels of the remaining six dimensions, an orthogonal main array plan was employed. These procedures resulted in the final selection of 960 choice tasks, with a total of 1920 health states. Two levels of randomization were employed for the DCE component of the survey. First, each respondent was randomly allocated 16 of the 960 choice sets without replacement. Second, for each choice task the ordering of the health states was randomized (i.e. which state was option A and which was option B). Randomization of the ordering intends to prevent any possible ordering bias emerging from the task presentation. The ordering of attributes was not altered, as this has been shown to increase the complexity of the choice tasks [30].

2.3 Survey Structure

The choice tasks were administered in an online survey.

The survey had seven components, presented in the following order: (1) introduction; (2) informed consent; (3) age and gender for quota sampling purposes; (4) EORTC-QLQ C30 questionnaire; (5) DCE valuation tasks and feedback module; (6) socio-demographic characteristics and clinical questions; and (7) other HRQoL measures, including the EQ-5D-5L [31] and Kessler K-10 mental health questionnaire [32].

The DCE valuation tasks presented responders with two health states, defined as 'Situation A' and 'Situation B' and described in terms of the QLU-C10D and time, and asked them to choose which of them they preferred. Based on the results of pilot data from the Australian valuation study, choice tasks were presented with the attributes that differed between situations A and B highlighted in yellow [17]. Figure 1 presents an example of the choice tasks for the DCE component of the survey. The DCE feedback module collected participants' perception of the clarity and difficulty of the valuation task, and enquired whether responders employed a specific response strategy.

The original survey was developed in English and translated into Spanish by professional translators. The translation procedure included forward and backward translations and feedback from an in-country support team.

2.4 Sample and Responders' Recruitment

The target sample for the online survey was 1000 Spanish adults representative of the general population. This sample size was based on the findings of the Australian QLU-C10D valuation study [18] and has been seen to be in the upper half of samples used for valuation studies of this kind [33]. Eligible participants were aged 18–80 years old. Quotas were used for gender and age, to ensure representativeness. Representativeness between the sample and the Spanish National population [34–36] was assessed using a two-sample test of proportions.

Participants were recruited during the period 4–29 July 2019 using an existing online panel provided by Toluna and administered by SurveyEngine®, two market research companies with specialist knowledge in DCEs. The panel comprised individuals who signed up to answer surveys in exchange for points that can be traded for goods. Respondents, as members of an online panel, received a web link to the survey, which they could access at their leisure. Reminders were not given, but recruitment remained open until the sampling frame was reached.

After agreeing to the informed consent, responders completed the survey at their own convenience.

Table 1 The QLU-C10D health state classification system, how it maps to the 13 component items from the QLQ-C30, and the duration attribute included in the discrete choice experiment (DCE) valuation survey [18]

Dimension	Level	Stem	Descriptor	QLQ-C30 item scores
Physical functioning ^{a,b}	1	You have...	No trouble taking a long walk outside of the house	Item 2 (long walk) = 1
	2		No trouble taking a short walk outside of the house, but at least a little trouble taking a long walk	Item 3 (short walk) = 1 AND Item 2 \geq 2
	3		A little trouble taking a short walk outside of the house, and at least a little trouble taking a long walk	Item 3 = 2 AND Item 2 \geq 2
	4		Quite a bit or very much trouble taking a short walk outside the house	Item 3 \geq 3 AND Item 2 \geq 2
Role functioning	1	You are limited in pursuing your work or other daily activities...	Not at all	Item 6 = 1
	2		A little	Item 6 = 2
	3		Quite a bit	Item 6 = 3
	4		Very much	Item 6 = 4
Social functioning ^{a,c}	1	Your physical condition or medical treatment interferes with your social or family life...	Not at all	Items 26 AND 27 = 1
	2		A little	Items 26 OR 27 = 2 ^c
	3		Quite a bit	Items 26 OR 27 = 3 ^c
	4		Very much	Items 26 OR 27 = 4 ^c
Emotional functioning	1	You feel depressed...	Not at all	Item 24 = 1
	2		A little	Item 24 = 2
	3		Quite a bit	Item 24 = 3
	4		Very much	Item 24 = 4
Pain	1	You have pain...	Not at all	Item 9 = 1
	2		A little	Item 9 = 2
	3		Quite a bit	Item 9 = 3
	4		Very much	Item 9 = 4
Fatigue	1	You feel tired...	Not at all	Item 18 = 1
	2		A little	Item 18 = 2
	3		Quite a bit	Item 18 = 3
	4		Very much	Item 18 = 4
Sleep	1	You have trouble sleeping...	Not at all	Item 11 = 1
	2		A little	Item 11 = 2
	3		Quite a bit	Item 11 = 3
	4		Very much	Item 11 = 4
Appetite	1	You lack appetite...	Not at all	Item 13 = 1
	2		A little	Item 13 = 2
	3		Quite a bit	Item 13 = 3
	4		Very much	Item 13 = 4
Nausea	1	You feel nauseated...	Not at all	Item 14 = 1
	2		A little	Item 14 = 2
	3		Quite a bit	Item 14 = 3
	4		Very much	Item 14 = 4
Bowel problems ^{a,c}	1	You...	Do not have constipation or diarrhoea at all	Items 16 AND 17 = 1
	2		Have a little constipation or diarrhoea	Items 16 OR 17 = 2 ^c
	3		Have constipation or diarrhoea quite a bit	Items 16 OR 17 = 3 ^c
	4		Have constipation or diarrhoea very much	Items 16 OR 17 = 4 ^c
Duration	1	You will live in this health state for...	1 year, and then die	Not applicable
	2		2 years, and then die	Not applicable
	3		5 years, and then die	Not applicable
	4		10 years, and then die	Not applicable

^aThree dimensions of the QLU-C10D each involve two QLQ-C30 items

^bThe physical functioning dimension includes 'long walk' and 'short walk' from the QLQ-C30; for the DCE, the levels are determined together, but were presented in the DCE survey separately, as shown in Fig. 1

^cFor social functioning and bowel problems, the QLU-C10D level is determined by the maximum value of the two component items

If you had to choose between these two health states, which would you pick?



Fig. 1 Example of the discrete choice experiment survey

2.5 Analysis

Data were analysed using STATA/MP 14 and SPSS 24. QALYs are traditionally anchored at zero for dead and one for full health, which ensures ratio scale properties [37] and zero condition assumption of the QALY model [38]. In order to set dead at 0, the functional form was specified with first-level interactions between the QLU-C10D dimensions levels and duration, as recommended by Bansback and colleagues [19]. Note that the experimental design allowed for all these interactions. This functional form allowed for the systematic component of the utility function to tend to 0 whenever the duration variable tended to 0.

Two different models were fitted to the data. The first model was a generalized estimating equation that adjusted

the possible intra-individual correlation generated by the panel structure of the data using the AR command. This model has been seen to produce similar results to the conditional logit with cluster sandwich estimator in previous studies [14]. In the model, the utility of option *j* (Situation A or Situation B) of choice set *s* 1 to 960 choice sets, for survey respondent *i*, was assumed to be

$$U_{isj} = \alpha \text{TIME}_{isj} + \beta X_{isj} \text{TIME}_{isj} + \epsilon_{isj}$$

where α was the utility associated with a life year, X_{isj} a vector of dummy variables representing the levels of the QLU-C10D health state presented in option *j* and β the vector of utility weights associated with each level in each dimension within X_{isj} , for each life year. The error term e_{isj} was assumed

to have a Gumbel distribution. Estimation of utility decrements was calculated by dividing each of the dimension levels' β terms by α . In the presence of non-monotonicity for the coefficients of different level severities for the same item, these levels were collapsed, as is commonly done (e.g. [28–30]).

The second model was a mixed logit model. This was used to investigate the impact of intra-individual variance (i.e. preference heterogeneity between individuals) on the analysis results. In the second model given by

$$U_{isj} = (\alpha + \gamma_i) \text{TIME}_{isj} + (\beta + \mu_i) X_{isj} \text{TIME}_{isj} + \varepsilon_{isj},$$

the additional parameters γ_i and μ_i represent individuals' specific differences from the individual mean preferences, which are given by α and β . Individuals' specific differences were assumed to follow a multivariate normal distribution. Two sensitivity analysis were performed. First, a conditional logit model was fitted to the data and compared with the generalized estimating equation model. Second, the tariff generated including the full sample was compared with the tariff generated when omitting responders who considered only a few dimensions.

3 Results

3.1 Characteristics of the Sample

Of the 1625 participants who were invited to the study, 17 (1%) declined to participate, 236 (14.5%) entered the web link but dropped out before completing the survey (e.g. did not give informed consent, quit before the beginning of the valuation component or did not complete the valuation component, etc.) and 362 (22.3%) were excluded as their respective quotas were already full. The remaining 1010 participants completed the study. The median time taken to complete the survey progressively decreased from 33 seconds in choice set 1–9 s in choice set 16. In line with previous studies, there was variability in the time taken per choice set between respondents [18, 25]. Appendix Figure 1 reports the median and percentile time stamps (i.e. median, fifth, twenty-fifth, seventy-fifth and ninety-fifth percentiles of time taken in seconds to complete each progressive choice set, i.e. first, second, third etc.) for the Spanish sample (see electronic supplementary material [ESM]).

Table 2 presents the sociodemographic characteristics of the included sample. The sample had an equal distribution between males ($n = 500$, 49.5%) and females ($n = 510$, 50.5%) and good representation of all age groups of the general population. Of the 1010 participants interviewed, 347 stated they had a chronic disease (34.4%). This was lower and statistically significantly different than the percentage of

people with chronic disease in the general Spanish population. Regarding education, the most common responses were a university education ($n = 461$, 45.7%), completed high school ($n = 319$, 31.6%) and completed only compulsory education ($n = 230$, 22.7%). This was statistically significantly different from the percentages found in the Spanish general population, where a higher number of compulsory education and a lower number of high school and university education responders were found.

3.2 DCE Feedback Module

Table 3 reports the results of the questionnaire regarding survey difficulty. As can be seen, most responders (58.2%, $n = 587$) found the survey easier or the same as other surveys. Of the 1010 participants, 235 considered the presentation of the health states as very clear (23.3%), 445 as clear (44.1%), 189 as neither clear nor unclear (18.7%), 106 as unclear (10.5%) and 35 as very unclear (3.5%). Choosing between health states was considered easy or very easy by 340 responders (33.7%), neither easy nor difficult by 302 (29.9%) and difficult or very difficult by 368 responders (36.4%). Most of the participants (31.7%, $n = 320$) explained that they considered only the aspects that differ between choice options (i.e. highlighted in yellow), 187 (18.5%) stated that they considered all the aspects and 228 (22.6%) stated that they considered most of the aspects.

3.3 Generalized Estimating Equation Utility Decrements (Non-imposed Monotonicity)

Table 4 reports the generalized estimating equation β coefficients and associated utility decrements (i.e. β/α) for the EORTC QLU-C10 without adjustments for inconsistencies. The significance of coefficients indicate that the attribute level had a statistically significant impact on the responders' choice. The sign of the coefficients indicates whether this impact was positive or negative.

As can be seen, for all attributes, level 4 was statistically significant, but level 2 and 3 of some attributes were not. More precisely, five attributes (i.e. physical functioning, fatigue, lack of appetite, nausea and bowel problems) registered negative and statistically significant coefficients for level 2, level 3 and level 4, four attributes for level 3 and level 4 only (i.e. role functioning, social functioning, emotional functioning and pain) and one attribute (i.e. sleep) for level 4 only.

The attribute resulting in the largest utility decrements was physical functioning, ranging between -0.090 of level 2 and -0.254 of level 4. The second largest utility decrements were associated with pain, that is, level 3 (-0.117) and level 4 (-0.174). Five attributes reported utility decrements < 0.1 for all their severity levels, and these were social

Table 2 Background characteristics of participants

Characteristic	Number	Sample percent- age	Population per- centage	<i>p</i> -value of difference ^a
Gender				
Male	500	49.5	49.4	0.977
Female	510	50.5	50.6	
Age (years)				
18–29	163	16.1	15.9	0.862
30–39	184	18.2	18.4	0.870
40–49	216	21.4	21.6	0.877
50–59	192	19.0	18.9	0.935
60–69	143	14.2	14.2	1.000
70–80	112	11.1	11.0	0.919
Chronic diseases				
Yes	347	34.4	41.1	< 0.000
No	663	65.6	58.1	
Education				
Compulsory	230	22.7	45.0	< 0.000
Higher secondary	319	31.6	22.0	< 0.000
University	461	45.7	33.0	< 0.000
EQ-5D-5L				
Mobility Level 1	809	80.1		
Mobility Level 2	141	14.0		
Mobility Level 3	41	4.0		
Mobility Level 4	15	1.5		
Mobility Level 5	4	0.4		
Selfcare Level 1	934	92.5		
Selfcare Level 2	44	4.3		
Selfcare Level 3	29	2.9		
Selfcare Level 4	2	0.2		
Selfcare Level 5	1	0.1		
Usual Activities Level 1	845	83.6		
Usual Activities Level 2	112	11.1		
Usual Activities Level 3	34	3.4		
Usual Activities Level 4	14	1.4		
Usual Activities Level 5	5	0.5		
Pain/Discomfort Level 1	564	55.8		
Pain/Discomfort Level 2	326	32.3		
Pain/Discomfort Level 3	91	9.0		
Pain/Discomfort Level 4	26	2.6		
Pain/Discomfort Level 5	3	0.3		
Anxiety/Depression Level 1	675	66.8		
Anxiety/Depression Level 2	230	22.8		
Anxiety/Depression Level 3	75	7.4		
Anxiety/Depression Level 4	23	2.3		
Anxiety/Depression Level 5	7	0.7		
Kessler K-10 score, mean (SD)	20.1	(7.5)		

Age and gender population percentages were obtained from the 2017 United Nations Statistics division [36]; Person reporting a chronic disease percentages were obtained from 2014 Eurostat data [34]; Population percentages for education levels were obtained from the OECD 2014 Education at a glance [35]

^aOne sample test of proportions. H0: Statistically significant difference in proportions

SD standard deviation

Table 3 Difficulty of survey

Question	Levels	Frequency	Percentage
Was the survey easier or harder than most surveys?	Easier	147	14.6
	The same	440	43.6
	Harder	391	38.7
	Could not say	32	3.2
Was the presentation of health states clear?	Very clear	235	23.3
	Clear	445	44.1
	Neither clear nor unclear	189	18.7
	Unclear	106	10.5
	Very unclear	35	3.5
How difficult it was to choose between pairs of health states?	Very easy	108	10.7
	Easy	232	23.0
	Neither easy or difficult	302	29.9
	Difficult	314	31.1
	Very difficult	54	5.3
Did you have a strategy for choosing between the health states?	Considered all aspects	187	18.5
	Considered most of the aspects	228	22.6
	Considered only aspects in yellow	320	31.7
	Considered only a few aspects	161	15.9
	Other	43	4.3
	Did not have a strategy	71	7.0

functioning, emotional functioning, fatigue, sleep and bowel problems. Among them, the attribute with the smallest utility decrement was sleep (i.e. only level 4 had a statistically significant decrement and this was -0.031).

Three of the EORTC QLU-C10 attributes were not monotonically decreasing (i.e. larger decrements associated with smaller levels of severity); these were social functioning, lack of appetite and nausea. For all three attributes, decrements were larger for level 3 than for level 4.

3.4 Generalized Estimating Equation Utility Decrements (Imposed Monotonicity)

Table 5 presents the generalized estimating equation β coefficients and associated utility decrements (i.e. β/α) for the EORTC QLU-C10 in presence of adjustments for monotonicity. Appendix Figure 2 presents a graph showing the utility decrements per dimension (see ESM). Once again, the significance of coefficients indicates that the attribute level had a significant impact on the responders' utility, while the sign indicates the direction of this impact.

Similarly to the analysis without monotonicity adjustment, coefficients of level 4 were statistically significant for all attributes, but only five of the ten attributes reported significant coefficients in all their levels. These were physical functioning, fatigue, lack of appetite, nausea and bowel problems. The attribute with the largest decrements was

physical functioning, followed by pain, while the one with the smallest decrements was sleep.

The collapsed utility decrements for the three attributes that were not monotonically decreasing in the unadjusted analysis were -0.087 for social functioning level 3 and level 4, -0.096 for nausea level 3 and level 4 and -0.050 for lack of appetite level 3 and level 4.

3.5 Mixed Logit Utility Decrements (Imposed Monotonicity)

Table 6 presents the mixed logit β coefficients and associated utility decrements (i.e. β/α) for the EORTC QLU-C10 in the presence of adjustments for monotonicity. The Akaike information criterion (AIC) was 15,594, the Bayesian information criterion (BIC) 16,013 and the log likelihood -7747 . Once again, the significance of coefficients indicate that the attribute level had a significant impact on the responders' utility, while the sign indicates the direction of this impact.

Similarly to the generalized estimating equation analysis with adjustment for monotonicity, coefficients of level 4 were statistically significant for all attributes. However, only four of the ten attributes reported significant coefficients in all their levels. These are physical functioning, fatigue, lack of appetite and nausea. The attribute with the largest decrement for level 4 was physical functioning, followed by pain, while the one with the smallest decrement was sleep.

Table 4 Generalized estimating equation QLU C10D utility decrements (not adjusted for non-monotonicities)

Attribute	Level ^a	Beta (β)	Standard errors	Mean/utility decrement
Time coefficient (α)	(linear)			0.562
Physical functioning	2	- 0.051**	0.021	- 0.090
Physical functioning	3	- 0.091**	0.022	- 0.163
Physical functioning	4	- 0.143**	0.019	- 0.254
Role functioning	2	- 0.002	0.017	- 0.003
Role functioning	3	- 0.059**	0.017	- 0.104
Role functioning	4	- 0.060**	0.015	- 0.107
Social functioning	2	-0.014	0.015	- 0.024
Social functioning	3	- 0.052**	0.015	- 0.093
Social functioning	4	- 0.048**	0.015	- 0.085
Emotional functioning	2	- 0.007	0.015	- 0.013
Emotional functioning	3	- 0.022*	0.016	- 0.039
Emotional functioning	4	- 0.044**	0.014	- 0.078
Pain	2	- 0.015	0.015	- 0.026
Pain	3	- 0.066**	0.017	- 0.117
Pain	4	- 0.098**	0.015	- 0.174
Fatigue	2	- 0.030**	0.014	- 0.053
Fatigue	3	- 0.034**	0.016	- 0.061
Fatigue	4	- 0.043**	0.014	- 0.077
Sleep disorders	2	0.003	0.015	+ 0.006
Sleep disorders	3	- 0.007	0.016	- 0.013
Sleep disorders	4	- 0.017*	0.013	- 0.031
Lack of appetite	2	- 0.022*	0.014	- 0.038
Lack of appetite	3	- 0.033**	0.015	- 0.059
Lack of appetite	4	- 0.025*	0.014	- 0.045
Nausea	2	- 0.037**	0.015	- 0.066
Nausea	3	- 0.059**	0.015	- 0.105
Nausea	4	- 0.051**	0.014	- 0.091
Bowel problems	2	- 0.023*	0.015	- 0.042
Bowel problems	3	- 0.047**	0.015	- 0.084
Bowel problems	4	- 0.049**	0.013	- 0.087

^aLevel 2 = a little; Level 3 = quite a bit; Level 4 = very much
* $p < 0.05$; ** $p < 0.01$

Table 5 Generalized estimating equation QLU C10D utility decrements (adjusted for non-monotonicities)

Attribute	Level ^a	Beta (β)	Standard errors	Mean/utility decrement
Time coefficient (α)	(linear)			0.560
Physical functioning	2	- 0.050**	0.021	- 0.089
Physical functioning	3	- 0.091**	0.022	- 0.162
Physical functioning	4	- 0.142**	0.019	- 0.254
Role functioning	2	- 0.002	0.017	- 0.003
Role functioning	3	- 0.058**	0.017	- 0.104
Role functioning	4	- 0.060**	0.015	- 0.107
Social functioning	2	- 0.013	0.015	- 0.023
Social functioning	3	- 0.049**	0.014	- 0.087
Social functioning	4	- 0.049**	0.014	- 0.087
Emotional functioning	2	- 0.007	0.015	- 0.013
Emotional functioning	3	- 0.021*	0.016	- 0.037
Emotional functioning	4	- 0.044**	0.014	- 0.078
Pain	2	- 0.015	0.015	- 0.027
Pain	3	- 0.066**	0.017	- 0.118
Pain	4	- 0.098**	0.015	- 0.175
Fatigue	2	- 0.030**	0.014	- 0.053
Fatigue	3	- 0.034**	0.016	- 0.061
Fatigue	4	- 0.043**	0.014	- 0.076
Sleep disorders	2	0.000	0.000	0.000
Sleep disorders	3	- 0.008	0.014	- 0.015
Sleep disorders	4	- 0.018*	0.012	- 0.033
Lack of appetite	2	- 0.020*	0.014	- 0.036
Lack of appetite	3	- 0.028**	0.013	- 0.050
Lack of appetite	4	- 0.028**	0.013	- 0.050
Nausea	2	- 0.036**	0.015	- 0.064
Nausea	3	- 0.054**	0.012	- 0.096
Nausea	4	- 0.054**	0.012	- 0.096
Bowel problems	2	- 0.024*	0.015	- 0.043
Bowel problems	3	- 0.047**	0.015	- 0.084
Bowel problems	4	- 0.049**	0.013	- 0.087

^aLevel 2 = a little; Level 3 = quite a bit; Level 4 = very much
* $p < 0.05$; ** $p < 0.01$

Utility decrements of the generalized estimating equation and the mixed logit were generally of the same size and in the same direction, highlighting a high degree of consistency

between the two analyses. The value range of the generalized estimating equation is slightly larger as evidenced by the generally larger utility decrements associated with level 4.

Table 6 Mixed logit QLU C10D utility decrements (adjusted for non-monotonicities)

Variable/domain	Level ^a	beta (β)	Standard errors	Standard deviation	Standard error	Mean/utility decrement
Time coefficient (α)	(linear)					1.182
Physical functioning	2	-0.137**	0.018	0.127**	0.021	-0.116
	3	-0.163**	0.020	0.158**	0.023	-0.138
	4	-0.216**	0.018	0.181**	0.019	-0.183
Role functioning	2	-0.013	0.017	0.139**	0.023	-0.011
	3	-0.108**	0.015	0.126**	0.025	-0.092
	4	-0.108**	0.015	0.126**	0.025	-0.092
Social functioning	2	0.0				0.000
	3	-0.092**	0.012	0.113**	0.021	-0.078
	4	-0.092**	0.012	0.113**	0.021	-0.078
Emotional functioning	2	-0.000	0.017	0.161**	0.024	0.000
	3	-0.032	0.017	0.108**	0.028	-0.027
	4	-0.113**	0.016	0.141**	0.020	-0.095
Pain	2	-0.004	0.017	0.138**	0.027	-0.006
	3	-0.125**	0.017	0.145**	0.024	-0.105
	4	-0.177**	0.016	0.247**	0.021	-0.149
Fatigue	2	-0.058**	0.016	0.092**	0.033	-0.049
	3	-0.070**	0.017	0.066*	0.034	-0.059
	4	-0.089**	0.016	0.144**	0.022	-0.076
Sleep disturbances	2	0				0
	3	0				0
	4	-0.045**	0.013	0.082*	0.028	-0.038
Appetite loss	2	-0.038*	0.015	0.077**	0.029	-0.032
	3	-0.052**	0.014	0.052**	0.018	-0.044
	4	-0.052**	0.014	0.052**	0.018	-0.044
Nausea	2	-0.028	0.016	-0.086*	0.045	-0.024
	3	-0.086**	0.017	0.114**	0.022	-0.073
	4	-0.099**	0.015	0.093**	0.028	-0.084
Bowel problems	2	-0.019	0.016	0.107**	0.027	-0.016
	3	-0.063**	0.017	0.122**	0.028	-0.053
	4	-0.065**	0.015	0.098**	0.026	-0.055

^aLevel 2 = a little; Level 3 = quite a bit; Level 4 = very much

* $p < 0.05$; ** $p < 0.01$

3.6 Sensitivity Analyses

Appendix table 1 in the ESM presents the results of the QLU C-10 utility decrements when estimated using the conditional logit model. As it can be seen, differences in utility decrements between the methods were ≤ 0.01 in absolute size and were not systematic. In the conditional logit model, role functioning level 4 was inconsistent, while this was not the case for the generalized estimating equation. The remaining inconsistencies occurred in both models.

The second sensitivity analysis tested the exclusion of responders who reported only focusing on a subset of dimensions. This sensitivity analysis had a non-systematic and

modest effect on the value set (results available from the authors on request).

3.7 QLU-C10D Health State Value Estimation

Utility decrements from the monotonically ordered generalized estimating equation model can be used for the calculation of the utility associated with different EORTC QLU-C10 health states. This is done by subtracting the utility decrements associated with each of the item levels from 1 (i.e. full health). For example, EORTC QLU-C10 state 2411111111 would be calculated as: $1 - 0.089 - 0.107 - 0 - 0 - 0 - 0 - 0 - 0 - 0 - 0 = 0.804$.

The utility associated with the PITS state (i.e. level 4 in all attributes), which indicates the worst possible attainable health, is -0.043 . This is similar to other international studies (e.g. [18, 25]). Of the 1,048,0256 described by the QLU-C10, 316 are worse than death based on the Spanish value set.

4 Discussion

This study collected preferences for a set of EORTC QLU-C10D health states to derive a value set representative of the Spanish adult population. For all attributes, statistically significant decrements were associated with the worst level (level 4), reflecting that people have a strong preference to not live with this degree of dysfunction or symptom severity. For five of the attributes, namely physical functioning, fatigue, lack of appetite, nausea and bowel problems, decrements were statistically significant for all levels, while for four attributes, namely role, social and emotional functioning and pain, decrements were statistically significant for level 3 and 4 but not for level 2. Lack of monotonicity occurred in three attributes, but this was adjusted through collapsing the two dimensions levels and re-estimating the model. The monotonically ordered generalized estimating equation model represents a consistent tariff that can be used for the derivation of health state values to use in economic evaluations as a second-best option when the preferred GPBM is not included in the trial of interest, or as a useful alternative to performing CUA sensitivity analyses.

The rank order of QLU-C10 items found in this study is in line with those of previous QLU-C10 valuations, with generic items generally receiving larger weights compared with cancer-specific items. Among the generic items, physical functioning reported the largest decrements followed by pain and role functioning. These findings are in line with those in Germany [14], Austria, Italy, Poland [26] and France [39]. Among the cancer-specific items, nausea reported the largest utility decrement followed by bowel problems, which mimicked the same rank order found in the Austrian valuation study [16].

In the current study, weights for three items did not decrease monotonically. Lack of monotonicity for some of the QLU-C10 items was observed also in other valuation studies, for example those of Austria, Australia and the Netherlands, as well as in numerous valuations of other PBMs (e.g. [40–43]). There are multiple possible explanations for the observed inconsistencies, such as, among the others, the size of the descriptive system (i.e. number of attributes) in relation to the sample of the study (i.e. number of participants), the wording of the attributes (i.e. responders' understanding of the descriptors and labels), the administration mode (i.e. engagement of participants in online

surveys) or the fact that those items do not have a stable impact on utility. While the current study accounted for the lack of monotonicity by collapsing levels 3 and 4, therefore generating a consistent tariff, the reasons underlining the observed inconsistencies are of interest and should be investigated in future research.

Two of the three cancer-specific items, fatigue and lack of appetite, were associated with small weights, a result that was in line with the attribute decrements of the QLU-C10 found in Australia and Germany [14, 18] and with preferences registered in other PBM experimental studies [44]. A possible explanation for this might be that members of the general public do not have experience of cancer symptoms, and how these severely impact the HRQoL of patients [45–47], or may have some experience of those symptoms, but at a lesser degree of intensity. Valuation studies eliciting values from cancer patients are currently ongoing. Results of these studies will inform on whether the small weights associated with cancer attributes depend on the population performing the valuation task or the comparative relevance of these attributes in explaining HRQoL against generic attributes.

In the current study, two different modelling approaches were employed, namely generalized estimating equation and mixed logit. The models reported similar mean utility decrements. We chose to estimate the QLU-C10 tariff using a generalized estimating equation, as in economic evaluation mean responses are more important than variability in preferences between different groups (i.e. preference heterogeneity) [18].

The current study developed a value set for the QLU-C10, a condition-specific PBM. Value sets for a number of other condition-specific PBMs exist, including those for the Amyotrophic Lateral Sclerosis utility index (ALS) [48], the Health Assessment Questionnaire (HAQ) [49], the Asthma Quality of Life Questionnaire (AQLQ) [50], the Exacerbation Utility (Exact U) [51], the Multi Sclerosis Impact Scale (MSIS) [49], the NEWQoL 8D [41] and the Dementia Quality of Life measure (DEMqoL) [42]. Valuation studies of condition-specific PBMs differ substantially in terms of the chosen preference elicitation technique (discrete choice experiment, time trade-off, standard gamble, rating scale etc.), the population providing preference weights (general public, patients, professionals etc.) and the country in which preference weights were obtained [52]. These choices have an impact on the relevance of the value sets for different application contexts. For example, a value set obtained from general public responders may be preferred over a value set obtained from patients by HTA bodies that take the payer perspective, while patients' index scores may be better suited to investigations of large patient registries, population health studies and for personalized medicine [e.g. 53]. The Spanish value set for

the QLU-C10 was developed to facilitate the conduct of economic evaluation in Spain, where condition-specific PBM values are accepted and used [6].

This study has some limitations. First, it used an online administration procedure. While this has the advantage of being cheaper and less time consuming than face-to-face administration, it may be associated with poorer quality of data due to responders' reduced engagement with the task and strategic behaviours [54–56]. To assess whether this occurred, time stamps were investigated, without finding any relevant problem. Yet, this possibility cannot be entirely ruled out. Second, the colour coding approach used a yellow colour for the non-overlapping items. Currently, there is a debate in the DCE literature with some supporting this approach, and others championing intensity colour coding as a better alternative [57]. While the approach adopted for task presentation is reasonable, it is important to acknowledge that other colour coding variants might have been used. Third, the sample reported higher education levels than the Spanish general population. From a theoretical point of view, this may have an impact on the values obtained, for example because of different understanding of the task, different priorities and different underlying health of responders with varying education levels. Yet, departures in the sample representativeness such as those reported in this study are similar to the ones of other valuation studies [58]. Moreover, there is evidence that educational levels do not substantially impact values when elicited through time trade-off [59, 60], albeit this evidence is not available for DCE. Fourth, in the DCE, the physical functioning domain was presented as two attributes (related to short walk and long walk) to simplify the wording of the attribute. However, the underlying design treated it as a single attribute. This may have had an impact on the values obtained, resulting in a larger decrement associated with physical functioning. Despite these limitations, this study also has important strengths. It generated the first value set for the QLU-C10 based on preferences elicited from the Spanish population. The availability of a value set for this population allows for a more theoretically sound alternative to the use of mapping techniques from non-PBMs to PBMs and from PBMs to non-PBMs [24]. It also allows easy access to values for end users, researchers and policy makers. Furthermore, by being part of a broader European programme of research, the current study allows comparative assessments of preferences for health states relevant to cancer populations across European countries. Finally, the methods and the DCE design employed in the current research have been previously tested in an experimental study [17] and have been already used for the estimation of QLU-C10 preference weights in different European countries, which increases the confidence in the results obtained.

5 Conclusion

This study generated a QLU-C10 tariff for Spain, which can be used in future economic evaluations or for comparative research across cancer populations in Europe.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40273-021-01058-x>.

Declarations

Funding Open access funding provided by Università Commerciale Luigi Bocconi within the CRUI-CARE Agreement. The study was funded with grant number 002-2014/European Organisation for Research and Treatment of Cancer.

Conflict of interest Aureliano Paolo Finch, Richard Norman and Rosalie Viney are members of the EuroQol Group. Eva Gamp, Bernahard Holzner, Madeleine King and Georg Kemmeler are members of the EORTC Quality of Life Group. The authors declare no conflicts of interest.

Ethical standards All procedures performed in this study were in accordance with the ethical standards of the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Ethics approval The study was approved by the Medical University of Innsbruck Ethics Committee, with approval number 20151207–1336.

Data availability statement Upon completion of the project, data will be made available for approved requests through the EORTC website at: <https://www.eortc.org/data-sharing/>.

Author contributions APF participated in the analysis of data, write up of the manuscript and revisions. EG, RV and BH participated in the writing of the manuscript and the revisions. RN and GK participated in the study preparation and data collection, analysis of the data, write up of the manuscript and revisions.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

1. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. *Health Econ.* 2018;27(1):7–22. <https://doi.org/10.1002/hec.3564>.

2. Brazier J, Ratcliffe J, Salomon J, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. 2nd ed. Oxford: Oxford University Press; 2016.
3. Fayers PM, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. Wiley; 2013.
4. W. en S. Ministerie van Volksgezondheid, 'Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidszorg - Publicatie - Zorginstituut Nederland', Feb. 29, 2016. <https://www.zorginstituutnederland.nl/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg>. Accessed 13 Jan 2020.
5. NICE. Guide to the methods of technology appraisal. 2013.
6. Lopez-Bastida J, Oliva J, Gisbert R, Mar J, Puig-Junoy J. Spanish recommendations on economic evaluation of health technologies. *European Journal of Health Economics*. 2010;11(5):513–20. <https://doi.org/10.1007/s10198-010-0244-4>
7. Finch AP, Brazier JE, Mukuria C. What is the evidence for the performance of generic preference-based measures? A systematic overview of reviews. *Eur J Health Econ*. 2018;19(4):557–70. <https://doi.org/10.1007/s10198-017-0902-x>.
8. Richardson J, Iezzi A, Khan MA. Why do multi-attribute utility instruments produce different utilities: the relative importance of the descriptive systems, scale and "micro-utility" effects. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil*. 2015;24(8):2045–53. <https://doi.org/10.1007/s11136-015-0926-6>.
9. Brazier J, Rowen D. NICE DSU technical support document 11: alternatives to EQ-5D for generating health state utility values. London: National Institute for Health and Care Excellence (NICE); 2011.
10. Brazier J, Ara R, Rowen D, Chevrou-Severac H. A review of generic preference-based measures for use in cost-effectiveness models. *Pharmacoeconomics*. 2017;35(1):21–31. <https://doi.org/10.1007/s40273-017-0545-x>.
11. Longworth L, 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 Winch Engl*. 2014;18(9):1–224. <https://doi.org/10.3310/hta18090>.
12. Qian X, Tan RL-Y, Chuang L-H, Luo N. Measurement properties of commonly used generic preference-based measures in East and South-East Asia: a systematic review. *Pharmacoeconomics*. 2019. <https://doi.org/10.1007/s40273-019-00854-w>.
13. Garau M, Shah K, Towse A. Assessment and appraisal of oncology medicines: does NICE's approach include all relevant elements? (What can be learnt from international HTA experiences? Report for the Pharmaceutical Oncology Initiative). OHE Consulting, London; 2009.
14. Kemmler G, et al. German value sets for the EORTC QLU-C10D, a cancer-specific utility instrument based on the EORTC QLQ-C30. *Qual Life Res*. 2019;28(12):3197–211. <https://doi.org/10.1007/s11136-019-02283-w>.
15. Rowen D, et al. Deriving a preference-based measure for cancer using the EORTC QLQ-C30'. *Value Health J Int Soc Pharmacoecon Outcomes Res*. 2011;14(5):721–31. <https://doi.org/10.1016/j.jval.2011.01.004>.
16. King MT, et al. QLU-C10D: a health state classification system for a multi-attribute utility measure based on the EORTC QLQ-C30. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil*. 2016;25(3):625–36. <https://doi.org/10.1007/s11136-015-1217-y>.
17. Norman R, et al. Using a discrete choice experiment to value the QLU-C10D: feasibility and sensitivity to presentation format. *Qual Life Res*. 2016;25(3):637–49. <https://doi.org/10.1007/s11136-015-1115-3>.
18. King MT, 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):225–38. <https://doi.org/10.1007/s40273-017-0582-5>.
19. Bansback N, Brazier J, Tsuchiya A, Anis A. Using a discrete choice experiment to estimate health state utility values. *J Health Econ*. 2012;31(1):306–18. <https://doi.org/10.1016/j.jhealeco.2011.11.004>.
20. Norman R, et al. Valuing SF-6D health states using a discrete choice experiment. *Med Decis Mak*. 2014;34(6):773–86. <https://doi.org/10.1177/0272989X13503499>.
21. Badia X, Roset M, Herdman M, Kind P. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. *Med Decis Mak Int J Soc Med Decis Mak*. 2001;21(1):7–16. <https://doi.org/10.1177/0272989X0102100102>.
22. Johnson JA, Luo N, Shaw JW, Kind P, Coons SJ. Valuations of EQ-5D health states: are the United States and United Kingdom different? *Med Care*. 2005;43(3):221–8. <https://doi.org/10.1097/00005650-200503000-00004>.
23. Karlsson JA, et al. National EQ-5D tariffs and quality-adjusted life-year estimation: comparison of UK, US and Danish utilities in south Swedish rheumatoid arthritis patients. *Ann Rheum Dis*. 2011;70(12):2163–6. <https://doi.org/10.1136/ard.2011.153437>.
24. McTaggart-Cowan H, Teckle P, Peacock S. Mapping utilities from cancer-specific health-related quality of life instruments: a review of the literature. *Expert Rev Pharmacoecon Outcomes Res*. 2013;13(6):753–65. <https://doi.org/10.1586/14737167.2013.850420>.
25. Norman R, et al. U.K. utility weights for the EORTC QLU-C10D. *Health Econ*. 2019;28(12):1385–401. <https://doi.org/10.1002/hec.3950>.
26. Gamber EM, et al. EORTC QLU-C10D value sets for Austria, Italy, and Poland. *Qual Life Res*. May 2020. <https://doi.org/10.1007/s11136-020-02536-z>.
27. Luo N. Preferences for quality of life in oncology: are cancer patients and the general population different? Abstract presented at the 26th ISOQOL annual conference, 2019.
28. Aaronson NK, et al. The European Organization for research and treatment of cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*. 1993;85(5):365–76. <https://doi.org/10.1093/jnci/85.5.365>.
29. Street DJ, Burgess L. The construction of optimal stated choice experiments: theory and methods. Wiley; 2007.
30. Logar I, Brouwer R, Campbell D. Does attribute order influence attribute-information processing in discrete choice experiments? *Resour Energy Econ*. 2020;60:101164. <https://doi.org/10.1016/j.reseneeco.2020.101164>.
31. Herdman M, 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>.
32. Kessler RC, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med*. 2002;32(6):959–76. <https://doi.org/10.1017/S0033291702006074>.
33. Mulhern B, Norman R, Street DJ, Viney R. One method, many methodological choices: a structured review of discrete-choice experiments for health state valuation. *Pharmacoeconomics*. 2019;37(1):29–43. <https://doi.org/10.1007/s40273-018-0714-6>.
34. Eurostat, 'Eurostat Data Explorer'. 2020. <https://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do>. Accessed 6 July 2020
35. OECD. Education at a glance 2014: indicators by chapter. 2014. <http://www.oecd.org/education/education-at-a-glance-2014-indicators-by-chapter.htm>. Accessed 10 July 2020.

36. UNdata. Population by age, sex and urban/rural residence. 2020. <http://data.un.org/Data.aspx?d=POP&f=tableCode%3a22>. Accessed 6 July 2020
37. Roudijk B, Donders ART, Stalmeier PFM. Setting dead at zero: applying scale properties to the QALY model. *Med Decis Mak*. 2018;38(6):627–34. <https://doi.org/10.1177/0272989X18765184>.
38. Miyamoto J, Wakker P, Bleichrodt H, Peters H. The zero-condition: a simplifying assumption in QALY measurement and multi-attribute utility. *Manag Sci*. 1998;44:839–49. <https://doi.org/10.1287/mnsc.44.6.839>.
39. Nerich V, et al. French value-set of the QLU-C10D, a cancer-specific utility measure derived from the QLQ-C30. *Appl Health Econ Health Policy*. 2020. <https://doi.org/10.1007/s40258-020-00598-1>.
40. 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).
41. Mulhern B, et al. The development of a QALY measure for epilepsy: NEWQOL-6D. *Epilepsy Behav*. 2012;24(1):36–43. <https://doi.org/10.1016/j.yebeh.2012.02.025>.
42. Rowen D, et al. Estimating preference-based single index measures for dementia using DEMQOL and DEMQOL-proxy. *Value Health*. 2012;15(2):346–56. <https://doi.org/10.1016/j.jval.2011.10.016>.
43. Stevens K. Valuation of the child health utility 9D index. *Pharmacoeconomics*. 2012;30(8):729–47. <https://doi.org/10.2165/11599120-000000000-00000>.
44. Finch AP, Brazier J, Mukuria C. Selecting bolt-on dimensions for the EQ-5D: testing the impact of hearing, sleep, cognition, energy, and relationships on preferences using pairwise choices. *Med Decis Mak*. 2021;41(1):89–99. <https://doi.org/10.1177/0272989X20969686>.
45. Cella D, Davis K, Breitbart W, Curt G, Fatigue Coalition. Cancer-related fatigue: prevalence of proposed diagnostic criteria in a United States sample of cancer survivors. *J Clin Oncol Off J Am Soc Clin Oncol* 2004;19(14):3385–3391. <https://doi.org/10.1200/JCO.2001.19.14.3385>.
46. Servaes P, Verhagen C, Bleijenberg G. Fatigue in cancer patients during and after treatment: prevalence, correlates and interventions. *Eur J Cancer*. 2002;38(1):27–43. [https://doi.org/10.1016/S0959-8049\(01\)00332-X](https://doi.org/10.1016/S0959-8049(01)00332-X).
47. Bower JE. Cancer-related fatigue: mechanisms, risk factors, and treatments. *Nat Rev Clin Oncol*. 2014;11(10):597–609. <https://doi.org/10.1038/nrclinonc.2014.127>.
48. Beusterien K, Leigh N, Jackson C, Miller R, Mayo K, Revicki D. Integrating preferences into health status assessment for amyotrophic lateral sclerosis: the ALS Utility Index. *Amyotroph Lateral Scler Mot Neuron Disord Off Publ World Fed Neurol Res Group Mot Neuron Dis*. 2005;6(3):169–76. <https://doi.org/10.1080/14660820410021339>.
49. Versteegh MM, Leunis A, Uyl-de Groot CA, Stolk EA. Condition-specific preference-based measures: benefit or burden? *Value Health*. 2012;15(3):504–13. <https://doi.org/10.1016/j.jval.2011.12.003>.
50. Young TA, Yang Y, Brazier JE, Tsuchiya A. The use of Rasch analysis in reducing a large condition-specific instrument for preference valuation: the case of moving from AQLQ to AQL-5D. *Med Decis Mak*. 2011;31(1):195–210. <https://doi.org/10.1177/0272989X10364846>.
51. Petrillo J, Cairns J. Development of the EXACT-U: a preference-based measure to report COPD exacerbation utilities. *Value Health*. 2011;14(4):546–54. <https://doi.org/10.1016/j.jval.2010.10.032>.
52. Rowen D, Brazier J, Ara R, Azzabi Zouraq I. The role of condition-specific preference-based measures in health technology assessment. *Pharmacoeconomics*. 2017;35(S1):33–41. <https://doi.org/10.1007/s40273-017-0546-9>.
53. Appleby J, Poteliakhoff E, Shah K, Devlin N. Using patient-reported outcome measures to estimate cost-effectiveness of hip replacements in English hospitals. *J R Soc Med*. 2013;106(8):323–31. <https://doi.org/10.1177/0141076813489678>.
54. Jiang R, Muehlbacher A, Shaw JW, Lee TA, Walton S, Pickard AS. Comparing online and face-to-face data quality and preferences in a health valuation study. *Patient Patient Centered Outcomes Res*. 2019;12(4):435–435.
55. Jiang R, et al. Comparison of preferences and data quality between discrete choice experiments conducted in online and face-to-face respondents. *Med Decis Mak Under Rev* (forthcoming).
56. Jiang R, et al. Comparison of online and face-to-face valuation of the EQ-5D-5L using composite time trade-off. *Qual Life Res*. 2021;30(5):1433–44. <https://doi.org/10.1007/s11136-020-02712-1>.
57. Jonker MF, Donkers B, de Bekker-Grob E, Stolk EA. Attribute level overlap (and color coding) can reduce task complexity, improve choice consistency, and decrease the dropout rate in discrete choice experiments. *Health Econ*. 2019;28(3):350–63. <https://doi.org/10.1002/hec.3846>.
58. Al Shabasy SA, Abbassi MM, Finch AP, Baines D, Farid SF. The EQ-5D-5L valuation study in Egypt. *Pharmacoeconomics*. 2021. <https://doi.org/10.1007/s40273-021-01002-z>.
59. Yang Z, van Busschbach J, Timman R, Janssen MF, Luo N. Logical inconsistencies in time trade-off valuation of EQ-5D-5L health states: whose fault is it? *PLOS One*. 2017;12(9):e0184883. <https://doi.org/10.1371/journal.pone.0184883> (set 2017).
60. Jin X, Liu GG, Luo N, Li H, Guan H, Xie F. Is bad living better than good death? Impact of demographic and cultural factors on health state preference. *Qual Life Res*. 2016;25(4):979–86. <https://doi.org/10.1007/s11136-015-1129-x>.