



# The societal monetary value of a QALY associated with EQ-5D-3L health gains

Laura Vallejo-Torres<sup>1,2,3</sup> · Borja García-Lorenzo<sup>2,4</sup> · Oliver Rivero-Arias<sup>2,3,5</sup> · José Luis Pinto-Prades<sup>6</sup>

Received: 3 December 2018 / Accepted: 14 November 2019 / Published online: 28 November 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

There is an extensive body of empirical research that focuses on the societal monetary value of a quality-adjusted life year (MVQALY). Many of these studies have found the estimates to be inversely associated with the size of the health gain, and thus not conforming to the linearity assumption imposed in the QALY model. In this study, we explore the extent to which the MVQALY varies when it is associated with different types and magnitudes of quality of life (QoL) improvements. To allow for a comprehensive assessment, we derive the MVQALY corresponding to the full spectrum of health gains defined by the EQ-5D-3L instrument. The analysis was based on a large and representative sample of the population in Spain. A discrete choice experiment and a time trade-off exercise were used to derive a value set for utilities, followed by a willingness to pay questionnaire. The data were jointly analysed using regression analyses and bootstrapping techniques. Our findings indicate that societal values for a QALY corresponding to different EQ-5D-3L health gains vary approximately between 10,000€ and 30,000€. MVQALY associated with larger improvements on QoL was found to be lower than that associated with moderate QoL gains. The potential sources of the observed non-constant MVQALY are discussed.

**Keywords** Willingness to pay · Quality-adjusted life year · Social perspective · EQ-5D

**JEL Classification** D61 · I10 · I18

## Introduction

There is a broad body of research on the societal monetary value of a quality-adjusted life year (QALY) [1, 2]. The motivation of these studies is, in most cases, to provide an estimation of the threshold value that would indicate the maximum cost per QALY for a technology to be considered cost effective. The underpinning idea is that health care funding decisions should be informed by the strengths of preferences of those members of society affected by these decisions [3]. Therefore, the cost-effectiveness threshold, upon which funding decisions are made, ought to reflect society's value for health gains.

It has been argued, though, that the monetary value of a QALY (MVQALY) is not a relevant source of information when allocating resources in the presence of fixed health care budgets [4]. The reason being that to adopt a new technology that imposes additional costs on the health care system, displacement of existing services is often required, resulting in health losses for individuals elsewhere. In such contexts, the threshold should represent the cost per QALY

---

✉ Laura Vallejo-Torres  
laura.vallejo@ulpgc.es

<sup>1</sup> Department of Quantitative Methods in Economics and Management, University of Las Palmas de Gran Canaria, Campus de Tafira, 35017 Las Palmas, Spain

<sup>2</sup> Fundación Canaria Instituto de Investigación Sanitaria de Canarias (FIISC), Canary Islands, Spain

<sup>3</sup> Health Services Research on Chronic Patients Network (REDISSEC), Canary Islands, Spain

<sup>4</sup> Assessment of Innovations and New Technologies Unit, Hospital Clínic, University of Barcelona, Barcelona, Catalonia, Spain

<sup>5</sup> National Perinatal Epidemiology Unit (NPEU), Nuffield Department of Population Health, University of Oxford, Old Road Campus, Headington, Oxford OX3 7LF, UK

<sup>6</sup> University of Navarra, Navarra, Spain

of displaced services (i.e. the opportunity cost), which allows the assessment of whether the health expected to be gained from the use of a new technology exceeds the health expected to be forgone elsewhere as other services are displaced.

Fixed budget constraints might characterise most, but not all decision-making contexts, and information on the MVQALY would still be relevant for, at least, two reasons. First, when displacement in health care services is not required, for instance, if funding is made available by raising new tax revenue or by decreasing allocation to other public sectors, the opportunity cost of investing in health care would fall across other alternative uses of public spending. A threshold reflecting the strengths of preferences of the public across different alternatives of consumption would arguably be more relevant in these cases [1, 5]. Secondly, using the opportunity cost approach alone might perpetuate the belief that the threshold reflects the marginal benefits of health care. Decision makers would not be made aware of technologies whose benefits measured according to society's views offset their costs but that might not be affordable under current budget constraints. Allowing for the identification and assessment of these circumstances could have implications for decisions on the size of the health care budget. Information on both the opportunity costs of health care funding decisions and society valuation of health gains is thus relevant to inform resource allocative decisions in health care [6].

On the societal value of health gains literature, the study of the maximum willingness to pay (WTP) has become the norm for the measurement of individuals' monetary valuation of a QALY. WTP for a QALY survey usually involves three steps: (1) to elicit in the QALY scale the utility associated with a health gain, (2) to elicit the WTP for that health gain, and (3) to combine the responses of these two estimates to arrive at a WTP for a QALY. Most of the empirical studies have consisted of asking individuals about their WTP for a small health gain, elicited in utility terms using standard gamble (SG) or time trade-off (TTO) techniques, and then aggregate this up to infer the WTP for a full QALY [7].

There are two main limitations with this general approach. These are the presence of non-traders and the linearity assumption in the valuation of health gains. These two issues are inherent to the traditional QALY model [8], but become even more problematic when combining information from the two dimensions involved in the WTP for a QALY framework, i.e. the substitution between quality of life and time/risk of death implied in the TTO/SG questionnaires, and the substitution between wealth and health implied in the WTP questionnaires. With regard to non-traders, the issue arises when individuals are willing to make a trade-off in one dimension but not in the other, which creates analytical difficulties. For instance, individuals reporting to be willing

to pay a non-zero quantity for an improvement on health for which they were not prepared to trade any time or risk of death yield an infinite MVQALY. On the linearity issue, the methods used so far implicitly assume that valuations of health gains are linear with respect to the size of the health gain. This assumption has been shown to not hold [9–13], which have led some to conclude that finding a single value of a QALY using these methods might not be possible [10]. While the source of this problem could somehow be related to limitations of the WTP techniques to capture the true value of health gains, it might also be the case that individuals experience diminishing marginal utilities for health gains and are thus unlikely to consider, say, a health gain twice as large as twice as valuable [14]. This has further led some authors to argue that it is inappropriate to express cost-effectiveness thresholds as maximum cost per QALY [15].

In this study, we aim to estimate the societal MVQALY in Spain, allowing for differences by types of health gains. In particular, we focus on the impact of using different magnitudes and dimensions of Quality of Life (QoL) improvements defined by all possible health states included in the EQ-5D-3L. This allows us to derive a range of societal monetary values for a QALY associated with all possible health profiles that relate to the most widely applied instrument used to describe health outputs in economic evaluations. Previous research has focused on exploring whether using different values for the duration and/or the severity of health states yields different results [9–11, 13]. To do so, these studies normally used two or more values for the duration of the health problem and/or two or more QoL profiles, selected a priori by the authors. In our study, we focus only on differences on QoL, but instead of using somehow arbitrary and relatively few values to define different QoL improvements, we estimate the monetary value of a QALY associated with the full spectrum of the EQ-5D-3L instrument. This provides a comprehensive assessment of the impact of using different QoL improvements on the corresponding MVQALY. We set the duration of the problem as fixed and equal to 1 month. The choice of such a short duration was informed by the findings of some of these previous studies that have emphasised that the size of the health problem shown to respondents ought to be small to partly address the issue that individuals reach their budget constraint when reporting WTP to avoid a health problem experienced for a long period [10].

To measure the monetary value of a QALY corresponding to every health state described by the EQ-5D-3L, we construct a *value set* for the EQ-5D-3L in terms of utilities and also in terms of WTP. By combining these two, we are able to compute the WTP for a QALY value corresponding to every possible EQ-5D-3L health state. This was accomplished by conducting a survey on a representative sample of the population in Spain and jointly modelling the responses

about their health preferences and their WTP values. We use a discrete choice experiment (DCE) rather than TTO or SG techniques to elicit health gains in utility terms. DCEs are particularly well suited to measure individual strengths of preferences, and have been recently applied to estimate the utilities associated with EQ-5D health states in the literature [16–19]. In DCE tasks, respondents are typically asked to select the option they would choose between two or more alternatives described in terms of a set of characteristics (attributes). In our study, the task implied individuals to indicate their preferred health state between two options described using EQ-5D-3L health profiles. This is an arguably simpler exercise than that required for TTO and SG techniques that involve going through an iterative process of identifying the point of indifference between two options [16]. It also avoids the issues related with respondents considering the health state shown to be worse than death, which requires further adjustments to the traditional TTO and SG methods. This simplicity provides us with the opportunity of conducting an online survey in a large and representative sample of the population that included the valuation of sufficient health states to create a complete value set plus a WTP questionnaire. Furthermore, this method deals with the non-trader issue discussed above, as DCEs simply ask individuals to choose the health state they preferred between two possibilities, rather than asking respondents to trade time or a risk of death to have a better quality of life. There is a limitation with the use of DCE designs though. The outputs produced by DCE data are in an arbitrarily scale, not anchored on the QALY utility scale of 0 (death) and 1 (perfect health). Therefore, as we discuss next, a rescaling approach was needed to produce values amenable to QALY calculations. We did so by including a single TTO task in the questionnaire that allowed the DCE values to be re-anchored in the QALY scale.

The methods we used are different to DCE designs that include cost as part of the attributes and estimate WTP values for the other attributes according to marginal rates of substitution. Such design would have precluded us from estimating a utility value set for the EQ-5D health states, and thus from expressing our results in terms of the corresponding WTP values for a QALY. Instead, we conducted a DCE using EQ-5D-3L dimensions as the only attributes, followed by a contingent valuation exercise to elicit WTP values related to these health states.

The rest of the paper is structured as follows. “[Methods](#)” describes the sampling and recruitment strategies, the DCE task, the WTP task, and the TTO task included in the questionnaire, as well as the modelling techniques used to analyse the data. Descriptive statistics of the sample and regression model results are provided in “[Results](#)”, alongside with the estimated ranges of the monetary value of a QALY. Finally, in “[Discussion](#)”, we summarise the main findings,

strengths and limitations of this study and provide some words of caution when interpreting some of our results.

## Methods

### Sampling and recruitment

The data were collected via an online survey. Respondents were recruited from an existing commercial internet panel with over 227,000 users in Spain. Participants were selected by quotas based on the demographic characteristics of the Spanish general population according to age (six categories: 18–24; 25–34; 35–44; 45–54; 55–64 and 65 or more), gender, and region of residence [Spain is divided into 17 areas called Autonomous Community (AC)].

Individuals who clicked on the invitation to the survey link were shown a project information sheet and then asked whether they consented to take part in the study. Following this, the quota questions were posed, and if the potential respondent belonged to a quota category already full was screened out of the study. Respondents entering the survey firstly completed the EQ-5D-3L for their own health and their self-reported health status. Then, each respondent answered the DCE task, the WTP task, and the TTO task (see “[Appendix 1: questionnaire \(selected sections\)](#)”). We used the EQ-5D-3L rather than the EQ-5D-5L version to keep the tasks faced by respondents as simple as possible. At the end of the survey, respondents were asked about their socioeconomic characteristics (income and employment status) and were given the opportunity to make comments on the questionnaire and indicate whether they found difficulties in completing the survey. Ethical approval for this study was granted by the Clinical Research Ethics Committee from Hospital Universitario Nuestra Señora de la Candelaria in the Canary Islands.

The questionnaire was piloted in a group of 15 researchers, clinicians and health economists, and in a sample of 200 respondents recruited from the internet panel. This resulted in minor improvements to the wording of the questionnaire being made and a change in the design of the DCE task as we detail below.

### DCE task

The DCE task consisted of pairs of scenarios based on the health states described by the five attributes of the EQ-5D-3L (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), with each attribute taking one of three possible levels of severity (none, some or severe problems—except for mobility that is defined as “confined to bed” at the highest severity level). Respondents were asked to imagine themselves living in two possible health states for a month, and to select which one they preferred.

The number of potential combinations of the EQ-5D-3L is 243 ( $3^5$ ). With two options to choose from in each choice scenario, this gives a possible 58,806 choices ( $243 \times 242$ ). To select the pairs of choices included in our choice set, we applied a fractional design by a means of a Bayesian efficient approach with informed priors based on the coefficients and standard errors estimated in a previous study in Spain [20].

The final choice set was reduced to 80 scenarios which were split into ten blocks (“Appendix 2: choice set”), so each respondent answered 8 DCE questions (the pilot included 10 DCE questions per individual which was found to be too long and tiring by pilot respondents). The choice set had a series of restrictions: dominated scenarios (i.e. when one health state is better in each of the five EQ-5D-3L domains) were excluded and, after the pilot, we decided to only allow three out of the five EQ-5D-3L domains to vary between the two health states shown to respondents. The reason was that participants of the pilot found the task too difficult when every domain could take a different value. The choice set design was programmed in R and had a D-error of 0.206. We checked the choice set comprised a broad variety of mild and severe health states. The order of appearance of the 8 DCE questions and, for each pair of health states, the state shown as state A (right on the screen) or state B (left on the screen) were randomly varied for each participant of the survey.

As noted above, a limitation of using the DCE approach is that the values produced by DCE designs are in an arbitrary scale that need to be anchored into the 0 (dead) to 1 (perfect health) scale. We considered including a time attribute as an additional domain into the DCE questionnaire, which has been done in previous studies to achieve this rescaling [16, 17, 21]. However, since our aim was to combine information on individuals’ valuation of health gains in terms of utility and in terms of WTP, we preferred to keep the duration of the health state constant, and short, in both parts of the questionnaire. Therefore, we used the duration of 1 month as part of the description of both health states rather than as a varying attribute. Other approaches to address rescaling include using death as a possible outcome in the choice set or as an opt-out choice. These have, though, been criticised for not conforming to the random utility theory underlying the DCE model, as a subset of respondents might consider all health states to be better than death [22]. An alternative solution is to use information on the TTO score of one health state from an external source (e.g. available EQ-5D-3L value set from the country of interest), or to undertake a TTO exercise within the same DCE sample to re-anchor their responses to the death-perfect health scale based on the preferences of the same individuals [23]. We decided to conduct the latter in this analysis and we compared the results when external TTO information was used.

## WTP task

After the DCE task, respondents were asked how much they would be willing to pay to avoid spending 1 month of their life with the health problems described by the EQ-5D-3L health states they faced at the DCE. However, to reduce the number of WTP questions posed to each respondent, the sample was randomly divided into two groups, with one half of the sample being asked their WTP to avoid the states shown as state A and the other half about states B. By doing so, each respondent answered 8 WTP questions and data were obtained for all the states used in the DCE.

WTP questions were posed as an out-of-pocket one-off payment that respondents would have to make to buy a medication that would cure the described health problems that would otherwise affect them during a month. The format of the possible WTP answers was six ranges covering a wide spectrum of quantities to limit framing bias: 0€–100€, 100€–500€, 500€–2000€, 2000€–6000€, 6000€–10,000€, or more than 10,000€, followed by an option that asked respondents to indicate the exact amount she/he would pay within the selected range, or a numerical open field that allowed participants to enter a value greater than 10,000€ if they chose the last option.

## TTO task

In this last task, respondents were asked to compare two possible hypothetical scenarios: in one scenario, they would live for 10 years with a described health problem and then die, and in the other they would live fewer years, starting at 5 years, in perfect health. Respondents had to choose their preferred scenario or indicate if they were indifferent. Depending on their answers, an iterative process was performed in which the time they would live without any health problem increased (decreased) by 6 months if the respondent chose the option with higher (lower) life expectancy. This process was carried out until the respondents indicated they were indifferent between the two options, or until their answers changed direction.

We used the health state corresponding to “22222” of the EQ-5D-3L algorithm, that is, the state with “some” problems in each of the five dimensions. This state allows us to assess a situation where all the dimensions of the EQ-5D-3L are affected. Furthermore, this state is thought to be not serious enough for individuals to consider it worse than death, which creates difficulties when eliciting TTO values, but not mild enough for a significant percentage of population to be unwilling to trade any time to improve their quality of life.



## Data analysis

Data collected in the DCE task were analysed using a conditional logit model. The data were analysed following the random utility theory framework that implies that the utility a person receives can be derived through a utility function with an explained component and a random component. The explained component is, in our case, a set of dummy variables corresponding to the levels for each EQ-5D-3L dimension. The model included ten coefficients corresponding to levels 2 and 3 of each dimension (MO = mobility, SC = self-care, UA = usual activities, PD = pain/discomfort, AD = anxiety/depression), using level 1 as the reference category. We conducted a main effect model, i.e. not including interactions in the analysis. This model took the form:

$$u_{ij} = \alpha_1 \text{MO}2_{ij} + \alpha_2 \text{MO}3_{ij} + \alpha_3 \text{SC}2_{ij} + \alpha_4 \text{SC}3_{ij} + \alpha_5 \text{UA}2_{ij} + \alpha_6 \text{UA}3_{ij} + \alpha_7 \text{PD}2_{ij} + \alpha_8 \text{PD}3_{ij} + \alpha_9 \text{AD}2_{ij} + \alpha_{10} \text{AD}3_{ij} + \varepsilon_{ij} \quad (1)$$

where  $u_{ij}$  is the estimated utility;  $i$  denotes the individual,  $j$  the DCE pair and  $\varepsilon_{ij}$  is the error term. We computed the mean TTO score of the 22222 health state obtained from our sample and rescaled the coefficients by dividing each  $\alpha$  coefficient by the value corresponding to the ratio:  $(22222_{\text{TTO}} - 1/22222_{\text{DCE}} - 1)$  [16]. In addition, we used data from the TTO score of the 22222 health state from the current Spanish EQ-5D-3L value set [24]. The rescaled coefficients can be interpreted as the disutility associated with a change in each level of each dimension of the EQ-5D-3L. To account for the fact that the disutility would be experimented for 1 month, as defined by the questionnaire, the rescaled coefficients were divided by 12 when computing the associated monetary value of a QALY.

Data collected in the WTP task were analysed using a log-transformed multilevel random effects model. Results obtained using fixed effects models were very similar. This model took the form:

$$\log(d_{ij}) = \gamma_1 \text{MO}2_{ij} + \gamma_2 \text{MO}3_{ij} + \gamma_3 \text{SC}2_{ij} + \gamma_4 \text{SC}3_{ij} + \gamma_5 \text{UA}2_{ij} + \gamma_6 \text{UA}3_{ij} + \gamma_7 \text{PD}2_{ij} + \gamma_8 \text{PD}3_{ij} + \gamma_9 \text{AD}2_{ij} + \gamma_{10} \text{AD}3_{ij} + \mu_i + \varepsilon_{ij} \quad (2)$$

where  $\log(d_{ij})$  is the logarithmic transformation of the reported WTP;  $i$  denotes the individual,  $j$  the EQ-5D state,  $\mu_i$  is the individual effect and  $\varepsilon_{ij}$  is the error term. The  $\gamma$  coefficients of the dummy variables were then re-transformed using the formulae  $(e^\gamma - 1) \times \text{mean}(\text{WTP})$  to compute the average WTP to avoid spending a month with a health problem corresponding to a change in each level of each dimension of the EQ-5D-3L.

We combined the information of each utility and WTP coefficient into a ratio that yields the WTP for a full QALY (i.e. the MVQALY) relative to each possible health gain described by the EQ-5D. By focusing on the estimated

coefficients across the sample, we are using the so-called aggregate approach, that is, we estimate the mean WTP and mean utility value across the full sample separately and we combine them into a ratio (ratio of means), as opposed to the disaggregated approach, which implies calculating this ratio for each individual and computing the mean across the sample (mean of ratios). MQALY corresponding to the move from level 2 or level 3 to level 1 in each domain was computed, and we also derived them for each of the possible moves from the 242 health states described by the EQ-5D-3L to perfect health.

Utility and WTP models were jointly estimated using bootstrapping techniques with 1000 replications, which allowed us to compute 95% confidence intervals around the WTP for a QALY estimate. We also conducted the models controlling for individual characteristics in terms of their self-assessed health, EQ-5D score, gender, age, and monthly income. We explored the existence of outliers, and we further analysed the data excluding respondents that: (1) indicated the same WTP for each one of the eight health states shown, and (2) spent less than 10 min to complete the survey. The average time required to complete the survey was estimated in 20 min. Subgroup analyses were also performed to compare the results among the sample with very good/good health and those reporting fair/bad/very bad health status, as well as among individuals with reported income below and above the sample median.

All analyses were undertaken using the software package Stata 14.0 (StataCorp USA).

## Results

### Descriptive statistics

Data were collected for 2003 individuals between December 2016 and January 2017. Demographic, self-assessed health and socioeconomic characteristics are presented in Table 1, as well as the degree of difficulty respondents found when answering the questionnaire.

Regional and gender representativeness was achieved in our sample, while the percentage of people aged 65 and over was lower (10% in our sample) than in the actual Spanish population [25] (22%). This is commonly the case when using online survey methods [26]. Responses to the self-assessed health were similar to the values collected by the National Health Survey in Spain (NHSS) in 2017 [27] which included over 22,000 adults. In our sample, 69% (66% in NHSS) reported very good or good health, 24% reported fair (24% in NHSS) and 7% (10% in NHSS) reported bad or very bad health. Monthly household income from respondents was also similar to that reported at NHSS: 58% reported a monthly household income lower than 2000€ (62% in

**Table 1** Sample characteristics

	<i>N</i>	%
Gender		
Male	966	48
Female	1037	52
Age		
18–24	159	7.9
25–34	321	16
35–44	510	26
45–54	482	24
55–65	332	17
65 or more	199	9.9
Region		
Andalucía	320	16
Aragón	61	3.1
Asturias	55	2.8
Canarias	95	4.7
Cantabria	28	1.4
Castilla y León	111	5.5
Castilla La Mancha	80	4
Cataluña	316	16
Valencia	225	11
Extremadura	47	2.4
Galicia	125	6.2
Balears	32	1.6
Rioja	9	0.5
Madrid	330	17
Murcia	54	2.7
Navarra	21	1.1
País Vasco	94	4.7
EQ-5D		
Mobility		
No problems	1700	85
Some problems	298	15
Confined to bed	5	0.3
Self-care		
No problems	1894	95
Some problems	101	5
Severe problems	8	0.4
Usual activities		
No problems	1691	84
Some problems	288	14
Severe problems	24	1.2
Pain/discomfort		
No problems	601	53
Some problems	498	44
Severe problems	30	2.7
Anxiety/depression		
No problems	1291	65
Some problems	640	32
Severe problems	72	3.6

**Table 1** (continued)

	<i>N</i>	%
General health		
Very bad	25	1.3
Bad	109	5.4
Fair	488	24
Good	1126	56
Very good	255	13
Monthly household income		
Less than 500€	90	4.5
500€–1000€	276	14
1000€–1500€	411	21
1500€–2000€	387	19
2000€–2500€	283	14.2
2500€–3000€	225	11
3000€–4000€	180	9
4000€–5000€	77	3.9
5000€–6.000€	28	1.4
6000€–7000€	11	0.6
7000€–9000€	7	0.4
More than 9000€	22	1.1
Occupation		
Permanent disability	62	3.1
Spouse's pension	10	0.5
Student	124	6.2
Retired	200	18
Employed	1049	53
Unemployed	318	16
Volunteering	5	0.3
Home care	134	6.7
Other occupation	44	2.2
Survey difficulty		
Not difficult at all	1170	58
Slightly difficult	504	25
Moderately difficult	288	14
Very difficult	40	2

NHSS), 26% between 2000€ and 3000€ (23% in NHSS) and 16% higher than 3000€ (15% in NHSS). 58% of respondents found the survey *not difficult at all*, while only 2% said it was *very difficult*.

### Regression models

Regression model results using the full sample are presented in Table 2. Four observations that reported a WTP higher than 1,000,000€ were considered outliers and were thus excluded from the analysis. The first two columns in Table 2 correspond to the results of the conditional logit model using data from the DCE. The coefficients were rescaled using the value of the TTO score of the 22222 state obtained from our

sample and estimated to be 0.564. This score was similar to the value previously estimated in the current value set of 0.572 [24]. There were no differences in the WTP for a QALY estimate when we use the TTO score estimated from our sample or the pre-existing tariff. Columns 3 and 4 report the coefficients related to the log-transformed WTP model and the corresponding re-transformed effects, respectively.

All coefficients had the expected sign and were statistically significant. In addition, coefficients related to level 3 had a greater impact, in absolute terms, than those related to level 2, both in terms of utility and WTP, as expected.

The corresponding MVQALY values derived by combining information on disutility (i.e. rescaled disutility coefficient divided by 12) and WTP are presented in column 5. These values varied between 9795€ for level 3 on the anxiety/depression domain to 25,503€ for level 2 on the mobility domain. In general, WTP for a QALY estimate was higher when they were associated with moderate health gains, i.e. to a change from “some problems” to “none”, than when corresponding to changes from “severe problems” to “none”. When focusing on moderate health problems (i.e. level 2), problems on mobility and self-care were associated with higher WTP for a QALY value, while when considering severe problems (i.e. level 3), health problems causing severe pain/discomfort had the highest estimated WTP for a QALY. 95% confidence intervals indicate that there is a degree of overlapping between domains/levels. The mean value across the 242 possible moves from an unperfect health state to perfect health was 14,104€, with a range varying from the aforementioned 9795€–25,503€ values.

Table 3 presents MVQALY estimates when we added controls for health status, gender, age and monthly household

income, and when we excluded respondents that reported the same WTP in all eight health states, and/or spent less than 10 min to complete the survey. Adding control variables slightly reduced the variability of the estimates produced by each domain/level. Excluding participants who reported same WTP values for each of the eight health states or took considerably less time than that expected to fill in the questionnaire, both resulted in higher WTP for a QALY estimate. This is particularly the case for the former, which implied excluding individuals who reported a WTP equal to 0€ in each WTP question. When we focus on the reduced sample (i.e. both not reporting same WTP and taking more than 10 min to complete the survey) and add control variables, the estimates varied from 11,145€ to 29,838€, with a mean value across the 242 potential health gains of 15,987€. We observe the same pattern in terms of (1) estimates related to moderate health problems (i.e. level 2 rather than 3) yield higher WTP for a QALY value (except for the pain/discomfort domain), and (2) the domains related to higher WTP values are mobility and self-care when the problems are moderate, and pain/discomfort when the problems are severe. Results from subgroup analyses indicated that the richest 50% had a higher associated WTP value for each level and dimension, except for moderate anxiety/depression, than the poorest 50%, while differences across the healthy/unhealthy subsamples did not show a clear pattern (Table 4).

## Discussion

In this study, we have explored the degree of variation in the societal MVQALY corresponding to different QoL improvements as described by all possible health states included in

**Table 2** Model results and WTP for a QALY estimate

Domain/level	Utilities		Willingness to pay (WTP)		
	Coeff. (1)	Rescaled coeff. (2)	Coeff. (3)	Re-transf. coeff. (4)	WTP for a QALY (4)/(2)/12 [95% CI] (5)
MO2	-0.2402***	-0.0691	0.1756***	146.86€	25,503€ [14,987–36,013]
MO3	-1.6388***	-0.4716	0.4165***	395.18€	10,056€ [7981–12,131]
SC2	-0.3297***	-0.0949	0.2056***	174.56€	22,078€ [15,450–28,706]
SC3	-1.3536***	-0.3895	0.4310***	412.05€	12,694€ [10,359–15,029]
UA2	-0.3431***	-0.0987	0.1898***	159.84€	19,426€ [13,397–25,454]
UA3	-1.3521***	-0.3891	0.4157***	394.26€	12,160€ [9706–14,614]
PD2	-0.3468***	-0.0998	0.1795***	150.40€	18,087€ [11,989–24,186]
PD3	-1.3889***	-0.3997	0.4929***	487.31€	14,631€ [11,646–17,616]
AD2	-0.2564***	-0.0738	0.1404***	115.29€	18,748€ [10,863–26,634]
AD3	-1.1037***	-0.3176	0.2919***	259.23€	9795€ [7155–12,435]
N	32,040 choices/~ 2003 participants		16,020 choices/~ 2003 participants		~ 2003 participants

Four observations with a WTP higher than 1,000,000€ were excluded

Coeff coefficient, CI confidence interval, WTP willingness to pay, QALY quality-adjusted life year

\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

**Table 3** WTP for a QALY estimate—sensitivity analyses

Domain/level	Exc. same WTP (1)	Exc. < 10 min (2)	Exc. same WTP and < 10 min [95% CI] (3)	Adding controls—full sample [95% CI] (4)	Adding controls—reduced sample [95% CI] (5)
MO2	30,637€	29,087€	34,298€ [15,990–52,606]	22,597€ [13,496–31,698]	29,838€ [13,233–46,443]
MO3	12,474€	9217€	10,906€ [8381–13,431]	10,299€ [8245–12,354]	11,145€ [8600–13,690]
SC2	27,294€	21,486€	25,669€ [16,861–34,477]	21,834€ [15,498–28,171]	25,282€ [16,779–33,786]
SC3	15,667€	11,365€	13,389€ [10,344–16,434]	13,148€ [10,629–15,667]	13,823€ [10,883–16,763]
UA2	22,882€	19,559€	21,954€ [13,793–30,115]	20,037€ [13,649–26,425]	22,578€ [14,539–30,617]
UA3	15,048€	11,689€	13,747€ [10,662–16,831]	12,605€ [10,056–15,153]	14,216€ [11,169–17,263]
PD2	20,859€	14,634€	16,530€ [10,156–22,904]	19,489€ [12,416–26,562]	17,590€ [10,030–25,150]
PD3	17,949€	13,695€	16,130€ [12,530–19,730]	15,335€ [12,200–18,469]	16,803€ [13,016–20,590]
AD2	22,357€	17,099€	19,433€ [9661–29,206]	19,257€ [11,198–27,315]	20,159€ [10,191–30,128]
AD3	12,301€	9599€	11,253€ [8185–14,322]	10,422€ [7711–13,132]	11,912€ [8683–15,140]
N	1788	1424	1287	1997	1284

(1) Excludes individuals reporting same WTP for all health states

(2) Excludes individuals who spent less than 10 min to complete the survey

(3) Exclude individuals reporting same WTP and taking less than 10 min to complete the survey (reduced sample)

(4) Includes health status, age, gender and income as control variables, in full sample

(5) Includes health status, age, gender and income as control variables, in reduced sample

the EQ-5D-3L instrument. Our findings indicate that societal values for a QALY related to different EQ-5D-3L health gains vary approximately between 10,000€ and 30,000€. MVQALY associated with larger improvements on QoL was found to be lower than that associated with moderate QoL gains, indicating that WTP is less than proportional to the size of the QoL improvement. Across the five EQ-5D domains, we find that individuals considered severe levels 4–7 times worse than moderate levels, but were only willing to pay between 2 and 3 times more to avoid severe rather than moderate health problems. As a result, the corresponding WTP for a QALY estimate is, in most cases, about twice

as large when we focus on moderate health gains. In fact, the highest (lowest) WTP for a QALY value is associated with the change in health individuals considered it to be the one causing the lowest (highest) disutility.

This apparent paradox is, however, a common finding in the literature [9, 10, 12, 28, 29], and the result of departures from the proportionality assumption and/or the lack of sensitivity of the WTP framework to elicit the true value of health gains. With regard to the latter, budgetary restrictions and insufficient adjustment bias have been hypothesised as potential reasons for the lack of proportionality in WTP responses [10]. We explored these possibilities within our

**Table 4** WTP for a QALY estimate—subgroup analyses

Domain/level	Among the richest 50% [95% CI] (1)	Among the poorest 50% [95% CI] (2)	Among the healthiest [95% CI] (3)	Among the sickest [95% CI] (4)
MO2	26,008€ [1133–50,884]	24,258€ [9166–39,350]	25,636€ [13,263–38,010]	25,905€ [6467–126,472]
MO3	12,922€ [9528–16,317]	8265€ [5924–10605]	11,176€ [8644–13,709]	7543€ [4236–10,850]
SC2	23,254€ [13,307–33,201]	21,102€ [11,843–30,361]	23,112€ [15,147–31,077]	18,858€ [5441–32,274]
SC3	15,116€ [11,227–19,006]	11,032€ [8170–13,894]	14,351€ [11,173–17,529]	9243€ [53,54–13,133]
UA2	21,201€ [10,104–32,299]	17,825€ [10,567–25,083]	19,405€ [12,100–26,709]	19,320€ [6627–32,013]
UA3	14,490€ [10,428–18,552]	10,520€ [7684–13,355]	12,788€ [9904–15,672]	10,379€ [5879–14,878]
PD2	24,876€ [13,858–35,893]	13,750€ [5734–21,767]	13,878€ [7686–20,070]	29,858€ [3706–56,011]
PD3	20,546€ [15,250–25,843]	11,036€ [7822–14,249]	14,418€ [11,050–17,786]	14,254€ [7937–20,571]
AD2	17,985€ [16,30–34,340]	18,307€ [8840–27,774]	18,126€ [8711–27,541]	19,977€ [7759–47,687]
AD3	11,147€ [6594–15,700]	8711€ [5781–11,641]	9593€ [6661–12,526]	9401€ [4680–14,121]
N	839	1164	1381	622

CI confidence interval



data. If budget constraints were an important source of non-linearities, we would expect individuals with higher incomes to show more proportional values. This was partly the case; WTP for a QALY value estimated among the richest 50% was, on average, 1.57 times larger for moderate than for severe problems, while there were, on average, 1.98 times larger among the poorest 50% (see Table 4). This suggests that budgetary constraints do play a role on the observed lack of linearity, but there is a degree of disproportionality that remains even among the richest individuals. We also explored the presence of anchoring and insufficient adjustment biases, which in our context are related to the fact that individuals tend to decide the minimum/maximum price they would pay to avoid a health problem and then they insufficiently adjust this price upward/downward according to the severity of the health state. We ranked the health states used in the questionnaire according to the disutility estimated by our analysis and created five equally sized groups (see Table 5). We computed the mean WTP reported for the health states included in each group, the mean disutility and the corresponding mean MVQALY. We observed that the largest gap on the associated WTP for a QALY value pertains to the difference between the group of least severe health states and the others, while the values were relatively proportional across the remaining four groups. This suggests that insufficient upward adjustment bias might be even more important than budgetary restrictions, as budget constraints should have a bigger impact on the group of more severe health states which are associated with higher WTP values.

Based on these findings, caution is needed when drawing conclusions on the results of this analysis. If taken as face value, the finding that moderate problems yield higher monetary values of a QALY than severe problems could be interpreted as implying that a lower cost per QALY threshold should be applied to technologies that achieve largest improvement of health than to technologies that achieve moderate improvements on health. However, this interpretation effectively assumes that the only reason behind the lack of proportionality observed in the WTP responses is that it reflects genuine preferences characterised by diminishing marginal utilities for health gains. However, our data suggest that the diminishing MVQALY is, at least partly, produced by the lack of sensitivity of WTP responses. We think this

is the most plausible explanation since the insensitivity of WTP to the size of benefit has also been observed in other areas such as WTP for risk reductions or environmental benefits [30, 31]. Therefore, the use of different decision thresholds based on this information is unlikely to be justified.

This study has a number of limitations. Online surveys provide a series of advantages but also difficulties in accessing individuals aged over 65 years old, and we observed responses which did not meet expected standards, such as the time expected to complete the survey. Methodologically, using the DCE approach we could only apply aggregate methods to combine responses on the utility and WTP values, which prevents us from the possibility of computing individual-level WTP for a QALY value. DCE techniques also implied the need to rescale the model outputs. In our study, we did so by conducting a TTO exercise, so the anchoring values were based on the elicited preferences from the same sample. However, the inclusion of this exercise is subjected to the aforementioned difficulties of TTO techniques. We aimed to limit these issues by only performing this exercise for one health state that was carefully chosen to attenuate the limitations that arise with TTO methods, which include individuals who are not willing to trade any time to avoid the health problem described or individuals who consider the state to be worse than death. Our data showed that 14% of the samples were not prepared to trade the minimum time that was allowed in the questionnaire (6 months) to avoid the problems described by the state 22222 (i.e. potential non-traders), and that about 5% of the sample were prepared to trade the maximum time that was shown (9.5 years) (i.e. potentially considering the state worse than death). A further advantage of the use of the health state 22222 to rescale the coefficients is that all the dimensions of the EQ-5D-3L are affected and thus the rescaling reflects the impact in each of the dimensions. For these reasons, the state 22222 might be considered the most suitable state to conduct the rescaling, but we acknowledge that the results are contingent to the state chosen. Using the estimated TTO value to conduct the rescaling also assumes that this mean score can be applied, on average, across the full sample. Finally, in this study we have focused on exploring the consequences of the lack of linearity when varying the size of the QoL gain, but we did not consider different

**Table 5** Mean values by level of severity

Severity level	Mean disutility [95% CI]	Mean WTP [95% CI]	Mean WTP for a QALY [95% CI]
Group 1	0.0300 [0.0253–0.0348]	609€ [451€–766€]	24,518€ [16,545€–32,491€]
Group 2	0.0510 [0.0494–0.0526]	637€ [557€–718€]	12,178€ [10,580€–13,777€]
Group 3	0.0669 [0.0651–0.0686]	733€ [640€–825€]	10,988€ [9579€–12,397]
Group 4	0.0791 [0.0775–0.0807]	769€ [702€–836€]	9735€ [8900€–10,571€]
Group 5	0.1018 [0.0969–0.1068]	1044€ [835€–1254€]	10,117€ [8285€–11,950€]

CI confidence interval

sources of variation in health outcomes, such as differences in durations of the health problems, or gains in life expectancy versus gains in quality of life. These have been the focus of previous research [9–13].

The monetary values of a QALY we estimate are similar to that provided in previous studies in Spain. Martín-Fernández et al. [32] estimated MVQALY ranging from 10,000€ to 28,000€ using a sample of 662 patients visiting health care facilities in Madrid. The EuroVaQ study run in nine countries with a sample of 2000€ individuals in Spain [26] estimated values in the range of 20,000€–40,000€ for this country. A previous study in Spain, which explored the sensitivity of the results to a series of variations in the design of the questionnaire, estimated values that varied between 5000€ and 124,000€ [10]. Our estimates are in line with recent evidence that have provided an estimation of the cost-effectiveness threshold based on the opportunity cost approach in Spain [6]. The average opportunity cost of health care funding decisions in Spain was proxied by the mean cost per QALY at the Spanish NHS and was measured between 22,000€ and 25,000€. This suggests that the thresholds derived from the opportunity cost framework are embodied within the range we have constructed using the societal valuation of health gains approach in Spain.

This study offers an estimation of the range of monetary values of a QALY associated with the full spectrum of

health gains defined by the EQ-5D-3L instrument. This was based on a large and representative sample of the population in Spain. The study shed light on the implications that the lack of proportionality has on the estimation of WTP for a QALY value and on the potential sources of the observed non-constant MVQALY. More work is needed before the observed variations can be considered to suggest that different decision thresholds according to the type and/or magnitude of the health gain might be appropriate.

**Acknowledgements** We want to thank Juan Manuel Ramos Goñi and Marcel Minke for their support with the survey design and implementation. We are very grateful to Juan Oliva Moreno, Fernando I. Sánchez Martínez and Eva M. Rodríguez Míguez for their comments on previous versions of this paper. We also want to thank all the researchers, health professionals and health economists that reviewed our survey, as well as all the participants that completed our questionnaire.

**Funding** This work was undertaken in the framework of activities run by the Network of Health Technology Assessment Agencies, funded by the Ministry of Health, Social Services and Equality in Spain.

### **Compliance with ethical standards**

**Conflict of interest** The authors declare that they have no conflict of interest.

## Appendix 1: Questionnaire (selected sections)

### Section A: Respondents' health status

<b>Section A</b>
<p>In this first section we will ask you some questions about your health today. Please, read each question carefully.</p> <p><b>Please, remind</b> this questionnaire is not an exam, therefore there are no correct or incorrect answers. It is important that you try to answer all questions. You do not need to show the answers to anyone as well as no one close to you will see your responses once you finish the questionnaire.</p>
<p>Indicate which statement best describes your health status TODAY</p> <p><b>Mobility</b></p> <p><i>Select one of the following options</i></p> <ul style="list-style-type: none"> <li><input type="radio"/> I have no problems in walking about</li> <li><input type="radio"/> I have some problems in walking about</li> <li><input type="radio"/> I am confined to bed</li> </ul>
<p><b>Self-care</b></p> <p><i>Select one of the following options</i></p> <ul style="list-style-type: none"> <li><input type="radio"/> I have no problems with self-care</li> <li><input type="radio"/> I have some problems washing or dressing myself</li> <li><input type="radio"/> I am unable to wash or dress myself</li> </ul>
<p><b>Usual Activities</b> (e.g. work, study, housework, family or leisure activities)</p> <p><i>Select one of the following options</i></p> <ul style="list-style-type: none"> <li><input type="radio"/> I have no problems with performing my usual activities</li> <li><input type="radio"/> I have some problems with performing my usual activities</li> <li><input type="radio"/> I am unable to perform my usual activities</li> </ul>
<p><b>Pain/discomfort</b></p> <p><i>Select one of the following options</i></p> <ul style="list-style-type: none"> <li><input type="radio"/> I have no pain or discomfort</li> <li><input type="radio"/> I have moderate pain or discomfort</li> <li><input type="radio"/> I have extreme pain or discomfort</li> </ul>
<p><b>Anxiety/depression</b></p> <p><i>Select one of the following options</i></p> <ul style="list-style-type: none"> <li><input type="radio"/> I am not anxious or depressed</li> <li><input type="radio"/> I am moderately anxious or depressed</li> <li><input type="radio"/> I am extremely anxious or depressed</li> </ul>
<p><b>During the last 12 months, would you say that your health has been very good, good, fair, poor or very poor?</b></p> <p><i>Select one of the following choices</i></p> <ul style="list-style-type: none"> <li><input type="radio"/> Very good</li> <li><input type="radio"/> Good</li> <li><input type="radio"/> Fair</li> <li><input type="radio"/> Poor</li> <li><input type="radio"/> Very poor</li> </ul>

## Section B: Discrete choice experiment

### Section B – Introduction

We are now going to ask you several questions where you should consider different **health states**. A **health state** is a health description about how you feel and how you manage in different situations of your day-to-day. Every health state presented below describes five health aspects.

Each question will show you two health states and you will be asked to indicate the health state that you consider to be better. Please, remember that there are no correct or incorrect answers.

Please, read the following example showing how to proceed through each question in this section.

#### Example

Below we show the description of two health states, health state A and health state B. Imagine that the next month of your life your health may be described as health state A or health state B. We ask you to indicate the health state in which you would prefer to live the next month of your life. If you consider that health state A is better than health state B, you should click on the box under health state A.

<b>Health state A</b>	<b>Health state B</b>
<i>I have no problems in walking about</i>	<i>I have no problems in walking about</i>
<i>I have no problems washing or dressing myself</i>	<i>I have no problems washing or dressing myself</i>
<i>I have no problems with performing my usual activities (e.g. work, study, housework, family or leisure activities)</i>	<i>I have no problems with performing my usual activities (e.g. work, study, housework, family or leisure activities)</i>
<i>I have no pain or discomfort</i>	<i>I have extreme pain or discomfort</i>
<i>I am extremely anxious or depressed</i>	<i>I am not anxious or depressed</i>
<input type="radio"/>	<input type="radio"/>

We will now ask you to indicate the health state you would prefer from the following 10 paired health states.

## Section C: Willingness to pay

### Section C – Willingness to pay

#### Section C - Introduction

In this section, we are going to ask you to think about the expenditure in which you sometimes need to incur to receive a medical treatment. For example, the National Health Service sometimes does not reimburse 100% of the cost of medicines we are prescribed, therefore we need to pay an amount of money at the pharmacy when buying such medicine. Also, sometimes we need to buy medical equipment (material for cures, prosthesis ...) that is not completely reimbursed. As a result, improving our health state usually costs us some money.

We will now ask you how much you would be **willing to pay for a medicine** that will improve your health.

We ask you to imagine that you may suffer a health problem and would spend the next month of your life in the health state indicated below. However, there is a medicine which can cure the condition, and if you take the medicine you will not suffer from that health problems on that month. To access this medicine, you would have to pay an amount of money only once. After that month, the health problem will be solved (with or without the medicine).

A usual issue regarding this type of questionnaires is that respondents are likely to state they would pay a higher price than they would actually do if they faced this situation in real life. We kindly ask you to consider this scenario as a real situation and think about the effect that amount paid for the medicine would have on your household budget and savings.

Please, read the following example showing how to proceed through each question in this section.

#### **Example:**

Imagine that your health the next month of your life may be described as the health state presented below.

How much would you pay to **avoid** spending one month in the health state below? Please, indicate a range that represents how much money you would pay at most.

I have no problems in walking about

I have no problems washing or dressing myself

I have no problems with performing my usual activities (e.g. work, study, housework, family or leisure activities)

I have no pain or discomfort

I am extremely anxious or depressed

Select one of the following choices

- 0€ - 100€
- 100€ - 500€
- 500€ - 2,000€
- 2,000€ - 6,000€
- 6,000€ - 10,000€
- More than 10,000€

Please, specify the exact amount you would pay within the range you have indicated

Click on and slide the bar to answer



Exact amount



## Section D: Time trade-off

### Section D - Introduction

We now ask you to compare **two hypothetical scenarios**. In one scenario, you would live for 10 years with a health problem detailed below as health state A and then die. In another scenario, you would live fewer years but without any health problem. We ask you to indicate what "life" you would prefer to live. If you consider that both scenarios presented below are equivalent (i.e., "you do not mind" in which of those two scenarios you would live) you may indicate "Indifferent".

Please, read the following example showing how to proceed through each question in this section.

#### **Example:**

Please, select which life, A or B, you think it is BETTER

<b><i>Life A: 10 years in Health State A</i></b>	<b><i>Life B: 1 year in full health</i></b>
<i>I have some problems in walking about</i>	<i>I have NO problems in walking about</i>
<i>I have some problems washing or dressing myself</i>	<i>I have NO problems with performing my usual activities</i>
<i>I have some problems with performing my usual activities</i>	<i>I have NO problems with performing my usual activities</i>
<i>I have moderate pain or discomfort</i>	<i>I have NO pain or discomfort</i>
<i>I am moderately anxious or depressed</i>	<i>I am NOT anxious or depressed</i>
<i>Followed by death</i>	<i>Followed by death</i>
<input type="radio"/>	<input type="radio"/>
<input type="radio"/> Indifferent	

**Appendix 2: Choice set**

Block	Pair		Block	Pair		Block	Pair	
1	21232	23221	4	13232	22222	7	23111	13213
1	13313	23211	4	11132	13111	7	32212	13222
1	32221	12322	4	33311	31323	7	31133	32122
1	12222	32112	4	32221	22232	8	13311	13132
1	23333	32332	4	11331	21211	8	21232	13332
1	22111	21231	5	21232	11333	8	31112	22312
1	21232	23131	5	23312	23133	8	32111	31221
1	31311	13312	5	22212	21233	8	21111	12131
2	23332	31312	5	13212	11322	8	32132	13332
2	11333	21132	5	22232	13332	8	32211	32133
2	31231	12331	5	22111	12213	8	31121	13123
2	22322	31222	5	11221	31111	9	32121	13123
2	23123	13233	5	32312	33211	9	32222	23232
2	31321	13322	6	12221	11323	9	31132	31211
2	21132	13232	6	21131	11233	9	32312	22333
2	32312	32223	6	23112	21213	9	23311	21332
3	23123	31113	6	32232	32311	9	11331	31221
3	12211	12123	6	13122	12232	9	13311	31111
3	31333	33113	6	23212	23123	10	23223	31221
3	22232	12333	6	22322	23212	10	32121	12331
3	33112	31213	6	13133	23111	10	22113	13133
3	23113	21323	7	23232	32222	10	32232	32321
3	22223	22312	7	22331	32221	10	23322	23133
3	21322	31221	7	22333	23223	10	23323	31123
4	23111	21122	7	12323	31322	10	22121	12231
4	33111	23221	7	33212	31223	10	31312	22313
4	11313	12112						

## References

- Ryen, L., Svensson, M.: The willingness to pay for a quality adjusted life year: a review of the empirical literature. *Health Econ.* (2014). <https://doi.org/10.1002/hec.3085>
- Nimdet, K., Chaiyakunapruk, N., Vichansavakul, K., Ngorsurachet, S.: A systematic review of studies eliciting willingness-to-pay per quality-adjusted life year: does it justify CE threshold? *PLoS One* **10**, e0122760 (2015). <https://doi.org/10.1371/journal.pone.0122760>
- Baker, R., Chilton, S., Donaldson, C., Jones-Lee, M., Lancsar, E., Mason, H., Metcalf, H., Pennington, M., Wildman, J.: Searchers vs surveyors in estimating the monetary value of a QALY: resolving a nasty dilemma for NICE. *Health Econ. Policy Law* **6**, 435–447 (2011). <https://doi.org/10.1017/S1744133111000181>
- Claxton, K., Martin, S., Soares, M., Rice, N., Spackman, E., Hinde, S., Devlin, N., Smith, P.C., Sculpher, M.: Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol. Assess.* **19**(1–503), 5–6 (2015). <https://doi.org/10.3310/hta19140>
- Vallejo-Torres, L., García-Lorenzo, B., Castilla, I., Valcárcel-Nazco, C., García-Pérez, L., Linertová, R., Polentinos-Castro, E., Serrano-Aguilar, P.: On the estimation of the cost-effectiveness threshold: why, what, how? *Value Heal.* (2016). <https://doi.org/10.1016/j.jval.2016.02.020>
- Vallejo-Torres, L., García-Lorenzo, B., Serrano-Aguilar, P.: Estimating a cost-effectiveness threshold for the Spanish NHS. *Heal. Econ.* (United Kingdom). (2018). <https://doi.org/10.1002/hec.3633>
- Vallejo-Torres, L., García-Lorenzo, B., Castilla, I., Valcárcel-Nazco, C., García-Pérez, L., Linertová, R., Polentinos-Castro, E., Serrano-Aguilar, P.: On the estimation of the cost-effectiveness threshold: why, what, how? *Value Heal.* **19**, 558–566 (2016). <https://doi.org/10.1016/j.jval.2016.02.020>
- Abellán-Perpiñán, J.-M., Pinto-Prades, J.-L., Méndez-Martínez, I., Badía-Llach, X.: Towards a better QALY model. *Health Econ.* **15**, 665–676 (2006). <https://doi.org/10.1002/hec.1095>
- Robinson, A., Gyrd-Hansen, D., Bacon, P., Baker, R., Pennington, M., Donaldson, C.: Estimating a WTP-based value of a QALY: the “chained” approach. *Soc. Sci. Med.* **92**, 92–104 (2013). <https://doi.org/10.1016/j.socscimed.2013.05.013>
- Pinto-Prades, J.L., Loomes, G., Brey, R.: Trying to estimate a monetary value for the QALY. *J. Health Econ.* **28**, 553–562 (2009). <https://doi.org/10.1016/j.jhealeco.2009.02.003>
- Bobinac, A., van Exel, N.J.A., Rutten, F.F.H., Brouwer, W.B.F.: GET MORE, PAY MORE? An elaborate test of construct validity of willingness to pay per QALY estimates obtained through contingent valuation. *J. Health Econ.* **31**, 158–168 (2012). <https://doi.org/10.1016/j.jhealeco.2011.09.004>
- Gyrd-Hansen, D., Kjaer, T.: Disentangling WTP per QALY data: different analytical approaches, different answers. *Health Econ.* **21**, 222–237 (2012). <https://doi.org/10.1002/hec.1709>
- Haninger, K., Hammitt, J.K.: Diminishing willingness to pay per quality-adjusted life year: valuing acute foodborne illness. *Risk Anal.* **31**, 1363–1380 (2011). <https://doi.org/10.1111/j.1539-6924.2011.01617.x>
- Nord, E., Enge, A.U., Gundersen, V.: QALYs: is the value of treatment proportional to the size of the health gain? *Health Econ.* **19**, 596–607 (2010). <https://doi.org/10.1002/hec.1497>
- Nord, E.: Beyond QALYs: multi-criteria based estimation of maximum willingness to pay for health technologies. *Eur. J. Heal. Econ.* **19**, 267–275 (2018). <https://doi.org/10.1007/s10198-017-0882-x>
- Bansback, N., Brazier, J., Tsuchiya, A., Anis, A.: Using a discrete choice experiment to estimate health state utility values. *J. Health Econ.* **31**, 306–318 (2012). <https://doi.org/10.1016/j.jhealeco.2011.11.004>
- Bansback, N., Hole, A.R., Mulhern, B., Tsuchiya, A.: Testing a discrete choice experiment including duration to value health states for large descriptive systems: addressing design and sampling issues. *Soc. Sci. Med.* **114**, 38–48 (2014). <https://doi.org/10.1016/j.socscimed.2014.05.026>
- Ramos-Goñi, J.M., Pinto-Prades, J.L., Oppe, M., Cabasés, J.M., Serrano-Aguilar, P., Rivero-Arias, O.: Valuation and modeling of EQ-5D-5L health states using a hybrid approach. *Med. Care* **00**, 1–8 (2014). <https://doi.org/10.1097/MLR.0000000000000283>
- Ramos-Goñi, J.M., Craig, B.M., Oppe, M., Ramallo-Fariña, Y., Pinto-Prades, J.L., Luo, N., Rivero-Arias, O.: Handling data quality issues to estimate the Spanish EQ-5D-5L value set using a hybrid interval regression approach. *Value Heal.* (2017). <https://doi.org/10.1016/j.jval.2017.10.023>
- Rivero-Arias, O., García-Lorenzo, B., Valcárcel-Nazco, C.: Comparing adolescents and adult preferences to EQ-5D-Y health states using profile case best-worst scaling. *Pap. Present. 33rd EuroQol Plenary Meet, Berlin.* 15–16 Sept. (2016)
- Norman, R., Viney, R., Brazier, J., Burgess, L., Cronin, P., King, M., Ratcliffe, J., Street, D.: Valuing SF-6D health states using a discrete choice experiment. *Med. Decis. Making* **34**, 773–786 (2013). <https://doi.org/10.1177/0272989X13503499>
- Flynn, T.N., Louviere, J.J., Marley, A.A., Coast, J., Peters, T.J.: Rescaling quality of life values from discrete choice experiments for use as QALYs: a cautionary tale. *Popul. Health Metr.* **6**, 6 (2008). <https://doi.org/10.1186/1478-7954-6-6>
- Norman, R., Mulhern, B., Viney, R.: The impact of different DCE-based approaches when anchoring utility scores. *Pharmacoeconomics.* **34**, 805–814 (2016). <https://doi.org/10.1007/s40273-016-0399-7>
- Badia, X., Roset, M., Monserrat, S., Herdman, M.: The Spanish VAS tariff based on valuation of EQ-5D health states from the general population. In: *EuroQol plenary meeting rotterdam*, pp 93–114 (1997)
- Instituto Nacional de Estadística (INE): Cifras de población. <https://www.ine.es> (2017). Accessed 26 Jan 2019
- Donaldson, C., Baker, R., Mason, H., Pennington, M., Bell, S., Lancsar, E., Jones-Lee, M., Wildman, J., Robinson, A., Bacon, P., Olsen, J.A., Gyrd-Hansen, D., Bech, T.K.M., Nielsen, J.S., Persson, U., Bergman, A., Protière, C., Moatti, J.P., Luchini, S., Prades, J.L.P., Mataria, A., Khatiba, R.A.: European value of a quality adjusted life year. [https://research.ncl.ac.uk/eurovaq/EuroVaQ\\_Final\\_Publishable\\_Report\\_and\\_Appendices.pdf](https://research.ncl.ac.uk/eurovaq/EuroVaQ_Final_Publishable_Report_and_Appendices.pdf) (2010). Accessed 13 Apr 2018
- Ministerio de Sanidad Consumo y Bienestar Social: Encuesta Nacional de Salud de España. <https://www.msbs.gob.es/estadEstudios/estadisticas/encuestaNacional/home.htm> (2017). Accessed 12 Dec 2017
- Gyrd-Hansen, D.: Willingness to pay for a QALY. *Health Econ.* **12**, 1049–1060 (2003). <https://doi.org/10.1002/hec.799>
- Hammitt, J.K., Haninger, K.: Willingness to pay for food safety: sensitivity to duration and severity of illness. *Am. J. Agric. Econ.* **89**, 1170–1175 (2007). <https://doi.org/10.1111/j.1467-8276.2007.01079.x>
- Hammit, J.K., Graham, J.D.: Willingness to pay for health protection: inadequate sensitivity to probability? *J. Risk Uncertain.* **8**, 33–62 (1999)
- Veisten, K., Hoen, H.F., Navrud, S., Strand, J.: Scope insensitivity in contingent valuation of complex environmental amenities. *J.*

- Environ. Manage **73**, 317–331 (2004). <https://doi.org/10.1016/j.jenvman.2004.07.008>
32. Martín-Fernández, J., Polentinos-Castro, E., del Cura-González, M.I., Ariza-Cardiel, G., Abaira, V., Gil-LaCruz, A.I., García-Pérez, S.: Willingness to pay for a quality-adjusted life year: an evaluation of attitudes towards risk and preferences. *BMC Health Serv. Res.* **14**, 287 (2014). <https://doi.org/10.1186/1472-6963-14-287>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.