Willingness to Pay for a Quality-Adjusted Life-Year: The Individual Perspective

Ana Bobinac, MSc, N. J. A. van Exel, MSc, Frans F. H. Rutten, PhD, Werner B. F. Brouwer, PhD

Department of Health Policy and Management and Institute for Medical Technology Assessment, Erasmus University Rotterdam, Rotterdam, The Netherlands

ABSTRACT

Objective: The aim of this study was to elicit the individual willingness to pay (WTP) for a quality-adjusted life-year (QALY).
Methods: In a Web-based questionnaire containing contingent valuation exercises, respondents valued health changes in five scenarios. In each scenario, the respondents first valued two health states on a visual analog scale (VAS) and expressed their WTP for avoiding a decline in health from the better health state to the worse, using a payment scale followed by a bounded open contingent valuation question.
Analysis: WTP per QALY was calculated for QALY gains calculated using VAS valuations, as well as the Dutch EQ-5D tariffs, the two steps in the WTP estimations and each scenario. Heterogeneity in WTP per QALY ratios was examined from the perspective of: 1) household income; and 2) the level of certainty in WTP indicated by respondents. Theoretical validity was analyzed using clustered multivariate regressions.

Results: A total of 1091 respondents, representative of the Dutch population, participated in the survey. Mean WTP per QALY was €12,900 based on VAS valuations, and €24,500 based on the Dutch EuroQoL tariffs. WTP per QALY was strongly associated with income, varying from €5000 in the lowest to €75,400 in the highest income group. Respondents indicating higher certainty exhibited marginally higher WTP. Regression analyses confirmed expected relations between WTP per QALY, income, and other personal characteristics.
Conclusion: Individual WTP per QALY values elicited in this study are similar to those found in comparable studies. The use of individual valuations in social decision-making deserves attention, however.
Keywords: cost-effectiveness, individual preferences, QALY, resource allocation, WTP.

Introduction

Decisions regarding reimbursement and allocation of funds within the health-care budget increasingly are influenced by the results of cost-effectiveness analysis (CEA). CEA evaluates two or more alternative interventions in terms of their benefits (expressed in a nonmonetary measure) and costs, and summarizes the result in an incremental cost-effectiveness ratio (ICER). The ICER represents the additional costs per additional health unit produced by one intervention in comparison to another. A common measure of health in this context is the quality-adjusted life-year (QALY), which comprises both length and quality of life. When using the QALY as outcome measure, the ICER represents the ratio of incremental costs per QALY gained. Typically, an intervention is considered cost-effective if the ICER falls below a certain cost-effectiveness "threshold," indicating some monetary value of a QALY. Some 10 years ago, Johannesson and Meltzer [1] argued that without explicating such a threshold value, CEA cannot be considered a proper decision-making tool, as it would lack a systematic and universally recognizable decision criterion.

Recent literature has seen a lively debate on implicit and explicit cost-effectiveness threshold(s), although without reaching consensus on the nature or height of an appropriate monetary value of a QALY [2–7]. In the mean time, various institutions and governmental bodies (such as the National Institute for Health and Clinical Excellence [NICE] in the UK, Swedish Pricing and Reimbursement board, Pharmaceutical Benefits Advisory Committee in Australia, CVZ in The Netherlands) have adopted

threshold values in the process of optimizing the allocation of health-care resources, albeit sometimes implicitly and inconsistently. The acceptable ranges of the monetary value of a QALY used in such decision-making, however, appear to be broad and tend to lack empirical underpinning [8,9]. This underlines the importance of further investigating the monetary value of a QALY.

The apparent reluctance to research and estimate a "true" value of a QALY has its roots in various arguments and in empirical, theoretical, and methodological challenges inherent to the process of obtaining such a number. For example, there is evidence that the willingness to pay (WTP) for a QALY is non-constant and dependent on the size, duration, and type of the health gain [10–15]. It might thus be impossible to elicit a unique individual WTP for a QALY, as suggested for example by Bleichrodt and Quiggin [16]. Matters are additionally complicated by the societal context of decision-making in health care. From the societal perspective, which aligns with the decision-maker's approach, the beneficiaries from health-care services need not be the payers of those services and therefore characteristics other than the size of the health gain may play a role in the valuation of QALYs. A review by Dolan et al. [17], for instance, showed the age of the beneficiary to be an important equity consideration that ought to be included in the social valuations of publicly provided health-care services. The discrepancy between individual and societal valuations, elicited from an ex ante or ex post perspective, could be considerable [18].

In spite of these problems, it is important to continue research in this area and work toward a higher level of transparency and consistency in societal decision-making. Seeking to find appropriate monetary values for QALY gains should not be seen as necessarily attempting to establish a firm link between CBA and CEA, but rather as an aid to decision-makers [12]. Indeed, Weinstein [19] recently concluded that "it is time to lay to rest

Address correspondence to: Ana Bobinac, Erasmus University Rotterdam (iBMG / iMTA), PO Box 1738, 3000 DR Rotterdam, The Netherlands. E-mail: a.bobinac@erasmusmc.nl

the mythical $50,000 per QALY standard and begin a real public discourse on processes for deciding what health care services are worth paying for."

This study aimed at eliciting the first empirical estimate of the monetary value of a QALY in The Netherlands. In doing so, it applies a carefully designed questionnaire, which draws on previous studies in this field. Specifically, it uses a contingent valuation approach, from the individual perspective and under certainty, to answer how much are Dutch citizens willing to pay for a QALY gain. This is one of three ways of determining what the optimal cost-effectiveness threshold should be [2]. First, the threshold can be inferred from previous decisions taken by leading institutions such as NICE [3,8]. Second, it can be set to exhaust an exogenously determined budget [6]. Third, it can be set by identifying the marginal value the society attaches to health. While WTP is a common way of deriving the value of a commodity, only a few studies have applied it in this area [15,20–22]. This study offers a more comprehensive approach to WTP elicitation, in terms of the number of health states valued (in absolute terms and per respondent; e.g., 20, 21, 15), ensuring a good coverage of the QALY scale. The two-step elicitation method WTP used in this study, using a payment scale (PS) followed by a bounded direct follow-up question, is also more comprehensive than was usually applied, because it combines two (linked) elicitation questions in order to arrive at a more precise estimate of the maximum WTP. This was done to combine the ease of a PS with the precision of an open-ended (OE) format. Moreover, throughout, the study applies several different ways of mitigating the hypothetical nature of the exercise. Finally, the robustness of the findings was ensured by sample properties and size, arguably leading to larger generalizability of the results.

Conversely, like previous studies, the current study employs the individual perspective to WTP elicitation. It is the first step in a larger research effort, designed to estimate the societal value of a QALY in The Netherlands. As a part of a larger study, our results offer a reference point for future findings and give important practical insight on how to derive the appropriate values to be used in social decision-making. We also aimed at comparing our findings with the empirical estimates reported in the literature.

In the following sections, we present the methodology and the design of the study. We then present and discuss the main results of the study and results of various subsample analysis. Finally, our aim was to compare the value of a QALY to already existing estimates, and discuss the underlying reasons for the differences we might find. Then, we discuss the practical implications of our findings.

Methods

WTP for a QALY was elicited in a representative sample of the general public in The Netherlands, by means of contingent valuation. Former research showed that the general population (i.e., a heterogeneous, less health-literate sample) elicits more certain, and less volatile health valuations and WTP estimates than patient and/or decision-maker groups [23]. The respondents were recruited by a professional Internet sampling company and the questionnaire administered in October 2008 through the Internet. Participants did not receive direct monetary compensation, but a small sum was donated to a charity of their choice, upon completion of the questionnaire.

Survey Instrument

In the introduction to the questionnaire, the respondents were briefed about the purpose and content of the questionnaire, and, to help them understand the WTP exercises, were offered three "warm-up" questions for nonhealth-related items (i.e., their WTP for: 1) a car; 2) housing; and 3) a pair of shoes). Next, the respondents were asked to describe their own health status using the EQ-5D profile and to rate own health, perfect health, and death on the EQ-VAS [21]. The respondents had the possibility to adapt the ratings until final confirmation was given.

After this introduction, the respondents solved five choice scenarios. Each scenario contained two EQ-5D health profiles or health states (please note, scenario design is discussed below). The respondents were asked which of the two health states they considered as the better one (see screen 1 in Appendix 1 found at: http://www.ispor.org/Publications/value/ViHsupplementary/ViH13i8_Bobinac.asp) and then requested to place the two health states on a visual analog scale (VAS) showing their previous valuations of current health, death, and perfect health (see screen 2 in Appendix 1). Next, the respondents were asked to imagine being in the health state they had chosen as the better one and to indicate their WTP to avoid spending 1 year in the health state they had chosen as the worse. This health loss (i.e., the difference between the better and the worse health state in the scenario) could be avoided by taking a painless medicine of unspecified properties once a month, for which one had to pay out-of-pocket in 12 monthly instalments (see Appendix 1 at: http://www.ispor.org/Publications/value/ViHsupplementary/ViH13i8_Bobinac.asp for the full question). The vehicle of health improvement was only described as "painless medicine" in order to remove any possible contamination of the health gain evaluation according to the means by which that improvement would be brought about [24].

Next, the WTP was elicited in a two-step procedure: first, a PS [25–29] was offered, followed by a bounded "OE" question. The boundaries in the "OE" question were determined by the amounts the respondents had indicated to certainly pay or certainly not pay in the PS phase.

In particular, in the first step, the respondents were presented with an ordered low-to-high PS of monthly installments (in €: 0, 10, 15, 25, 50, 75, 100, 125, 150, 250, 300, 500, 750, 1000, 1500, 2500), and asked to indicate the maximum amount they would certainly pay (screen 3) and the first amount they would certainly not pay (screen 4 [27]). By asking the respondents to identify all the amounts they would certainly pay and those that they would certainly not pay, the method provided information about the range of values over which people are uncertain [30]. In the second step, the respondents were presented with a bounded direct "OE" follow-up question and asked to indicate the maximum amount they would pay if asked to do so right now. This maximum WTP was deemed as the appropriate estimate to be used in the calculation of the WTP for a QALY, and is our central WTP estimate. This estimate was bounded by the higher and the lower value the respondents previously chose on the PS (screen 5). The combination of two WTP questions, although in the context of a bidding game, was applied before (e.g., [31,32]). The two-step contingent valuation approach was applied to arrive at a directly and precisely indicated estimate of the maximum WTP within a range of WTP which was informed by the results from the less precise, but informative and easy-to-use PS. This two-step approach also added information and potentially robustness to our findings because the respondents used two different valuation techniques within one questionnaire. The benefit of employing two different WTP formats, although in a context of two entirely separate WTP questions, was investigated by Johnson et al. [33].

Attention was also given to reducing the hypothetical bias inherent in contingent valuation exercises, through ex ante and

ex post mitigation [34]. Ex ante, the respondents were reminded to take their household income into consideration when solving the exercise [35]. Moreover, the visual image of health states rated on the VAS remained present on the right-hand side of the screen, as a reminder of the size of the health gain being valued (see Appendix 1 at: http://www.ispor.org/Publications/value/ViHsupplementary/ViH13i8_Bobinac.asp). Ex post, the respondents were asked on which element of household spending they would economize in order to be able to pay for the painless medicine (answer options were: 1) food; 2) clothing; 3) entertainment; 4) sport; 5) savings; 6) charity; and 7) other) [36]. To avoid respondent fatigue and repetition, this was asked only at the end of the first of the five scenarios. Finally, the respondents were asked to indicate the level of certainty in the answer provided. They were asked to imagine having to pay the stated amount in reality, and immediately, and the options included: 1) totally sure I would pay the stated amount; 2) pretty sure I would pay the stated amount; 3) neither sure nor unsure I would pay the stated amount; 4) not very sure I would pay the stated amount; or 5) unsure I would pay the stated amount. This follow-up question was introduced to identify a subset of responses whose valuations may more closely reflect their "true" WTP [36–38]. Nevertheless, being surer in the valuation does not necessarily imply that the elicited value is "true" or that it necessarily reflects the revealed preference. It is only assumed that the stated WTP will probably deviate more from "true" WTP when the respondents are less sure about their answers.

When the respondents chose €0 as their maximum WTP, they were asked to indicate the reason behind this preference (answer options were: 1) I am unable to pay more than €0; 2) avoiding the worse health state and remaining in the better health state in not worth more than €0 to me; 3) I am not willing to pay out of ethical considerations; 4) something else [with open text field for explanation]; options 1) and 2) were considered as true WTP, options 3) and 4) as a protest answer).

The scenarios were presented in a random order to the respondents as to control for possible order bias, although such effects may not be entirely possible to eradicate [39]. Still, by adopting a randomized order, the potential bias was distributed more or less evenly across the blocks.

Following the main part of the questionnaire, the respondents were asked about their socioeconomic and demographic characteristics.

The questionnaire was pilot tested in a random sample of 100 respondents in order to determine the plausibility and clarity of the tasks, the feasibility of the questionnaire as a whole, and to test the range of the PS. The respondents had several opportunities to express their opinion about the tasks at hand. The results of the pilot showed that the questionnaire was clear and feasible, with no evidence to support the claim that the task was found unrealistic. Moreover, the two-step contingent valuation exercise proved feasible. The results of the pilot did point out that the distribution and spread of the PS were not optimal; the initial scale encompassed three value categories above €2500 (i.e., €5000, €7500, and €10,000), which were never chosen. To avoid loss of information and possible anchoring to exaggerated high values, the maximum was set at €2500 for the main study and additional value categories were added to the scale around the most frequently chosen values.

Design of Scenarios

Forty-two health states were paired into 29 choice scenarios (see Appendix 2 at: http://www.ispor.org/Publications/value/ViHsupplementary/ViH13i8_Bobinac.asp) representing a fair spread of QALY gains across the utility plane (see Fig. 1). The majority of the pairs was originally applied for deriving the UK tariffs for the EQ-5D [40], and 16 out of the 29 pairs were also applied in deriving the Dutch tariffs [41]. The few scenarios that were not applied in deriving the UK or the Dutch tariffs were chosen for the purpose of testing other hypotheses, on which the current study does not focus. The 29 scenarios were split into 10 blocks of five scenarios, and randomly assigned to a bit more than 100 respondents per block. Two scenarios per block were
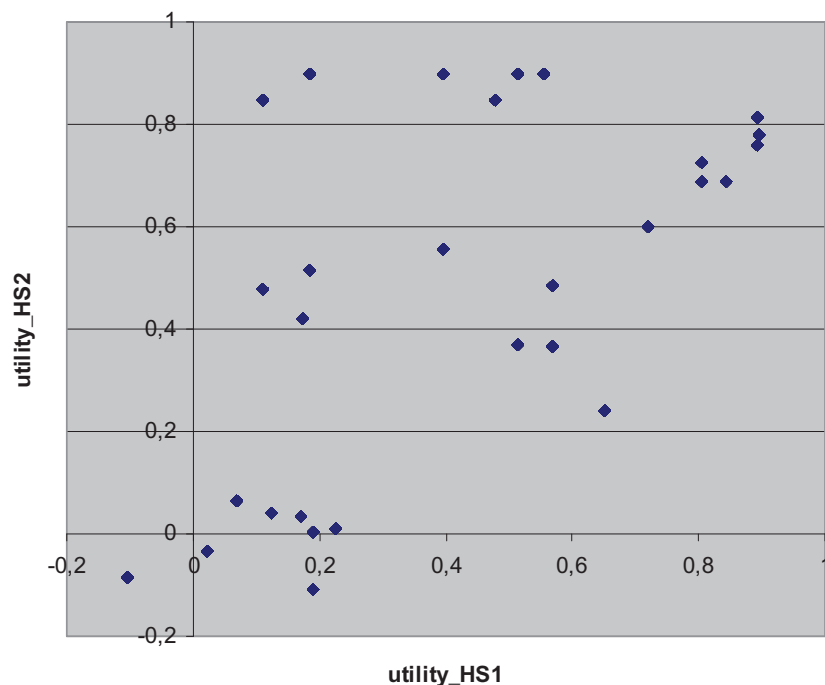

Figure 1 Spread of gains across the utility plain.

randomly assigned to one of the 10 blocks, and three were purposefully selected into blocks. These scenarios were assigned to blocks in order to ensure that in each block, the respondents encountered health gains situated on the low, middle, and high end of the utility scale (according to Dutch EQ-5D tariffs). Given the design, the changes in health between two health states (according to Dutch EQ-5D tariffs) ranged from 0.004 to 0.738 QALY, with a mean of 0.32 and a median of 0.34 QALY. Several scenarios were designed such that one health state was unambiguously better than the other.

QALY gains were also calculated from sample-specific VAS scores [28] obtained from the valuations in the questionnaire (i.e., "raw" scores); mean (and median) scores of perfect health and death were used for rescaling, based on the formula:

$$X_{rescaled} = \frac{X_{raw} - X_{mean\ of\ death}}{X_{mean\ of\ perfect\ health} - X_{mean\ of\ death}}$$

Combining the highest (€2500) and the lowest values (€10) of the PS with the minimum (0.004) and the maximum (0.738) QALY gains defined by the design produces an implicit maximum WTP of €7,500,000 (2500/0.004 * 12) and an implicit minimum WTP of €163 (10/0.738 * 12), with an implicit average WTP for a QALY of €17,862 (476.31/0.32 * 12) (see Appendix 2 at: http://www.ispor.org/Publications/value/ViHsupplementary/ViH13i8_Bobinac.asp). The ratio is multiplied by 12 because we ask about the monthly installment and a yearly health gain.

Analysis

WTP per QALY was calculated as the ratio of the WTP for avoiding the move from the better to the worse health state to the QALY difference between the two health states. This ratio was calculated for two utility elicitation techniques (i.e., using EQ-5D tariffs and EQ-VAS scores), two WTP elicitation techniques (i.e., PS and "bounded" OE formats), and for each of the five scenarios (i.e., taking the means of ratios of each individual scenario). The approach of taking the mean of ratios accounts for the individual variation in the marginal utility of income, and overall heterogeneity in preferences, because individuals' WTP for a QALY is directly imputed into the calculation of the mean. The most relevant WTP per QALY estimate was calculated based on valuations from the bounded direct "OE" follow-up question.

The heterogeneity in WTP per QALY ratios was primarily examined from the perspective of: 1) the level of household income, using the income categories presented in Table 1; and 2) the level of certainty in the WTP answers, by comparing the sample average WTP per QALY to the WTP per QALY of the respondents indicating the highest levels of certainty (pretty sure and totally sure).

The theoretical validity of our results was examined with a log-linear clustered multivariate regression analysis with raw WTP estimates and WTP per QALY estimates as dependent variables. Both variables were expected to increase with the level of household income and to decrease with the number of people depending on this income, while raw WTP was also expected to increase with the size of the projected health gain. Within the multivariate regression context, we also explored if the health status of respondents and/or existence of chronic illnesses would in any way affect the WTP per QALY estimate.

Separate regressions were conducted for each of the utility and WTP elicitation techniques. Variables and their associations were compared using parametric and nonparametric tests. The results of the PS were tested for order bias by comparing the WTP estimates between samples that solved the same blocks of scenarios in different orders. The specification of the PS, and the

Table 1 Summary statistics (n = 1091)

Variable	Mean	SD	Min	Max
Age	42.1	12.1	18	65
Sex (% male)	0.47	0.50		
Marital status:				
Married (% yes)	0.61	0.49		
Divorced (% yes)	0.10	0.31		
Single (% yes)	0.24	0.43		
Widow (% yes)	0.03	0.16		
Unknown (% yes)	0.02	0.14		
Children (% yes)	0.56	0.50		
Number of children (n = 3070)	2.23	10.1	1	10
Income groups:				
Group 1 (% <€1000)	0.13	0.33		
Group 2 (% >€999 and <€2000)	0.34	0.48		
Group 3 (% >€1999 and <€3500)	0.40	0.49		
Group 4 (% >€3499)	0.12	0.33		
Number of people living on household income	2.44	10.4	1	20
University education (% yes)	0.36	0.48		
Employment status				
Employed (% yes)	0.62	0.48		
Unemployed (% yes)	0.17	0.38		
Student (% yes)	0.06	0.25		
Housewife/husband or retired (% yes)	0.14	0.35		
Health status				
EQ-5D (Dutch tariff)	0.84	0.22	−0.26	1
EQ-VAS	78.5	170.1	0.00	100
Suffering a chronic illness (% yes)	0.39	0.94		
Completion time of the questionnaire	18.8	60.13	9	61

mid-point and end-point bias, were investigated by examining response patterns on the PS, both in the pilot and the main study. The relationship between EQ-VAS and EQ5D tariffs was checked for consistency. All analysis was conducted using STATA for Windows version 10.

Results

One thousand ninety-one respondents, representative of the Dutch population according to age (18–65 years), sex, and education, participated in the survey. The description of the sample is given in Table 1. The respondents were predominantly married, employed, and in very good health (EQ-5D 0.84; EQ-VAS 78.5) (39% of the sample reported suffering from a chronic condition, and although the severity of the condition was not specified, given the average score on the EQ-VAS and EQ-5D tariff, we can assume that the respondents predominately suffered from very mild or mild chronic conditions.). The sample average net household income of €2564 a month, with an average of 2.44 household members depending on that income, adequately represents the Dutch national figures for 2008 [42].

WTP for Nonhealth Items

The respondents gave plausible estimates of WTP for a car (mean €10,900, median €7000), a pair of shoes (mean €109, median €80), and housing (mean €201,600, median €200,000) of their choice. From this, we inferred that the respondents understood the exercise, although the focal point of the exercise—health—may be more difficult to value, as normal (direct) market prices are absent.

Utilities

The correlation between utility scores obtained from the EQ-VAS scores and the Dutch EQ-5D tariffs was low (r = 0.24). The

average health gain was 0.32 (SD 0.2; median 0.34) based on EQ-5D tariffs and 0.33 (SD 0.29; median 0.25) based on the EQ-VAS. Although the average scores do not differ considerably, there is a statistically significant difference between them ($P = 0.02$). The tests of consistency between EQ-VAS and the Dutch EQ-5D tariffs conducted on the level of particular health states showed that the two valuation techniques especially provided similar valuations for health states situated in the middle range of the utility scale. It was tested and confirmed that the better health states received, on average, higher valuations on the EQ-VAS. The respondents reversed the ranking (i.e., valuing the obviously worse health state higher on the EQ-VAS) on average in only 7% of scenarios.

Patterns in WTP Answers

Data inspection did not disclose any unusual patterns. Less than 1.5% of the respondents chose the highest level offered on the PS (i.e., €2500). Sixty-two respondents indicated, in one or more scenarios, that they would not pay more than €0 for a health gain (only 23 respondents indicated €0 in all five scenarios). No consistent relationship was found between the size of the health gain, household income, and zero WTP. The interpretations of zero WTP were uniformly distributed among the offered explanations, and protest answers were observed in only 1.4% of all scenarios. We therefore proceeded with the analysis without specifically considering (or excluding) these responses.

The distribution of the certainty in the provided answers revealed that the majority of respondents (56%) was either pretty or totally sure that they would actually pay the stated amount for the specified health gain; 33% indicated uncertainty, 8% was not very sure they would pay, and 3% indicated they were unsure they would pay. The majority of the respondents would give up charitable donations or savings if they needed to pay for the medicine out-of-pocket.

Test results did not indicate that the mid-point or range bias played a noteworthy role in our study. The results show a highly left-skewed distribution of values chosen from the PS. Although the results did not concentrate at one particular amount on the scale, the majority of values chosen on the scale fell between €50 and €200. Finally, the tests showed that WTP per QALY elicited when the scenario offering the largest gain was presented first in a block did not differ from WTP per QALY estimates elicited when the scenario offering the smallest gain was presented first, thus refuting the order bias.

Maximum WTP per QALY

The estimates of WTP per QALY varied considerably with the method of calculation. Table 2 provides the breakdown of WTP per QALY values according to: 1) the source of health state valuations; 2) the two steps in the WTP elicitation (i.e., lower bound of the PS, that is, the amount people definitely would pay, and the OE follow up); and 3) the level of certainty. In the

bounded OE follow-up question, the respondents elicited a maximum WTP per QALY of €24,500. Estimates were systematically lower when QALY gains were calculated using the EQ-VAS scores (i.e., €12,900 and €17,000, rescaled on mean or the median). The estimates were higher among the respondents, indicating a high level of certainty in their answers, although the differences were not considerable. All estimates presented in Table 2 were statistically different from each other ($P = 0.00$). Finally, as an additional test, a zero WTP was assigned to responses that were "unsure" about the elicited WTP. This did not result in a significant change in the mean WTP ($P = 0.07$).

Subgroup Analyses

WTP per QALY varied considerably with household income in the expected direction, and reached €55,900 in the highest income group (Table 3). As noted before, the respondents indicating a higher level of certainty in their answers produced somewhat higher WTP per QALY estimates; those in the highest income group and with the higher certainty level elicited a mean individual WTP per QALY of €75,400 (using Dutch EQ-5D tariffs; see (3) in Table 3). VAS scores yielded considerably lower estimates (up to €35,300 for those in the highest income and certainty group; see (2) in Table 3). Differences in WTP per QALY were, however, only statistically significant between income group 4 and other groups ($P = 0.00$).

What Would You Not Pay?

On the PS, the respondents indicated that the minimum amount they were certainly not willing to pay for a QALY was €43,160 (see (3) in Table 4). Again, estimates were higher for respondents who were more certain in their answer (up to €48,600), for higher income groups (up to €86,100), and these characteristics combined (up to €114,900). Using VAS scores in the calculation of the amount the respondents were not willing to pay for a particular gain yields the estimate of up to €54,000 (see (2) in Table 4).

Theoretical Validity

Table 5 presents the results of multivariate logarithmic regressions with raw WTP values and WTP per QALY estimates as dependent variables. *t* Tests showed that all independent variables were statistically significant (at 1 or 5% level), and *F* test showed that the regression equations were statistically significant at any regular level. Results for health gains computed using median-rescaled VAS scores were omitted because they are highly comparable to the mean-rescaled ones.

The results are in line with a priori expectations; the WTP were positively associated with the size of the health gain and with household income, the latter effect being the strongest (beta coefficients, not presented here). Regression analysis of WTP per QALY estimates showed a positive association with the size of

Table 2 Willingness to pay (WTP) per quality-adjusted life-year (€, rounded to hundreds)

		All respondents: average [n = 1,091, f = 5,253]; (SD)	Certainty level: pretty sure or totally sure [n = 761, f = 2,984]; (SD)
1. WTP: EQ-VAS, mean rescaled	WTP: PS	9,600 (35,800)	10,400 (32,900)
	WTP: OE	12,900 (48,100)	13,100 (37,900)
2. WTP: EQ-VAS, median rescaled	WTP: PS	12,600 (47,100)	13,700 (43,200)
	WTP: OE	17,000 (63,200)	17,300 (49,800)
3. WTP: Dutch EQ-5D tariffs	WTP: PS	17,900 (172,100)	21,200 (181,600)
	WTP: OE	24,500 (213,600)	26,800 (204,300)

Table 3 Willingness to pay (WTP) per quality-adjusted life-year in different income groups and levels of certainty (€, rounded to hundreds)

		Income groups; (SD)				Income groups and certainty level: pretty sure or totally sure; (SD)			
		1 [n = 139, f = 672]	2 [n = 371, f = 1,806]	3 [n = 440, f = 2,117]	4 [n = 134, f = 658]	1 [n = 86, f = 318]	2 [n = 248, f = 964]	3 [n = 317, f = 1,262]	4 [n = 107, f = 440]
1. WTP: EQ-VAS, mean rescaled	WTP :PS	5,000 (12,900)	8,200 (41,100)	8,800 (22,100)	20,800 (60,900)	5,000 (7,600)	8,500 (30,100)	9,100 (19,600)	22,200 (63,600)
	WTP: OE	8,000 (27,300)	11,400 (63,500)	11,900 (28,500)	25,200 (62,000)	6,700 (10,500)	11,100 (42,100)	11,500 (22,400)	26,900 (64,200)
2. WTP: EQ-VAS, median rescaled	WTP: PS	6,500 (17,000)	10,800 (54,000)	11,500 (29,100)	27,300 (80,000)	6,600 (9,900)	11,100 (39,600)	12,000 (25,700)	29,100 (83,600)
	WTP: OE	10,500 (36,000)	15,000 (83,500)	15,700 (37,500)	33,200 (81,600)	8,800 (13,900)	14,700 (55,300)	15,100 (29,400)	35,300 (84,400)
3. WTP: Dutch EQ-5D tariffs	WTP: PS	8,200 (34,100)	14,300 (178,300)	15,100 (90,700)	47,100 (349,200)	8,000 (29,600)	11,400 (47,900)	17,400 (111,900)	63,900 (426,200)
	WTP: OE	12,600 (287,600)	18,000 (182,400)	21,100 (128,400)	55,900 (369,000)	11,100 (41,300)	15,100 (60,000)	22,900 (156,300)	75,400 (450,000)

Table 4 Willingness to pay (WTP) per quality-adjusted life-year upper bound, average, and different income groups and levels of certainty (€, rounded to hundreds)

	All respondents: average [n = 1,091, f = 5,253]; (SD)	Certainty level: pretty/totally sure [n = 761, f = 2,984]; (SD)	Income groups; (SD)				Income groups and certainty level: pretty/totally sure; (SD)			
			1 [n = 139, f = 672]	2 [n = 371, f = 1,806]	3 [n = 440, f = 2,117]	4 [n = 134, f = 658]	1 [n = 86, f = 318]	2 [n = 248, f = 964]	3 [n = 317, f = 1,262]	4 [n = 107, f = 440]
1. WTP: EQ-VAS, mean rescaled	23,800 (75,800)	22,500 (56,200)	18,000 (92,300)	22,400 (85,000)	22,100 (57,300)	39,150 (81,200)	12,500 (21,100)	20,400 (63,800)	20,000 (42,900)	41,000 (80,700)
2. WTP: EQ-VAS, median rescaled	31,300 (99,700)	29,500 (73,900)	23,700 (121,400)	29,400 (111,700)	29,100 (75,300)	51,500 (106,800)	16,400 (27,700)	26,800 (83,800)	26,300 (56,400)	54,000 (106,100)
3. WTP: Dutch EQ-5D tariffs	43,160 (308,500)	48,600 (351,000)	45,800 (412,900)	33,000 (264,400)	37,800 (210,000)	86,100 (502,900)	44,400 (421,600)	32,700 (265,000)	39,000 (248,900)	114,900 (612,700)

Table 5 Multivariate clustered regression analysis

	EQ-VAS scores (mean rescaled)						Dutch EQ-5D tariffs					
	WTP: PS			WTP: OE			WTP: PS			WTP: OE		
	Coef	SE	P > |t|	Coef	SE	P > |t|	Coef	SE	P > |t|	Coef	SE	P > |t|
DV: log(raw WTP)												
Heath gain: log(EQ-VAS)	0.05	0.02	0.06	0.06	0.03	0.02						
Health gain: EQ-5D tariff							0.13	0.05	0.02	0.13	0.06	0.04
Log(income)	0.74	0.07	0.00	0.81	0.09	0.00	0.73	0.07	0.00	0.79	0.09	0.00
Log(age)	−0.21	0.1	0.03	−0.33	0.11	0.00	−0.21	0.09	0.02	−0.33	0.11	0.00
Higher education	0.25	0.07	0.00	0.23	0.07	0.00	0.25	0.06	0.00	0.22	0.07	0.00
Number of people living on household income	−0.07	0.02	0.00	−0.09	0.03	0.00	−0.07	0.02	0.00	−0.08	0.03	0.00
Intercept	−0.55	0.57	0.34	−0.42	0.67	0.53	−0.58	0.56	0.29	−0.37	0.67	0.58
N		4982			4841			5029			5184	
R^2		0.11			0.12			0.12			0.18	
DV: log(WTP/QALY)												
Log(income)	0.71	0.08	0.00	0.78	0.10	0.00	0.72	0.07	0.00	0.79	0.10	0.00
Log(age)	−0.24	0.11	0.00	−0.35	0.12	0.00	−0.28	0.10	0.01	−0.40	0.11	0.00
Higher education	0.31	0.08	0.00	0.29	0.08	0.00	0.20	0.07	0.00	0.18	0.08	0.02
Number of people living on household income	−0.07	0.03	0.01	−0.08	0.03	0.01	−0.07	0.03	0.01	−0.08	0.03	0.01
Intercept	3,6	0.65	0.00	3,7	0.73	0.00	3,7	0.59	0.00	3,8	0.70	0.00
N		4841			4982			5029			5184	
R^2		0.07			0.08			0.06			0.06	

DV, dependant variable.

the household income as well, and a negative association with the number of household members supported by the household income.

In both sets of regressions, dependent variables increased with the level of education, as also found by Zethraeus [43], and decreased with age [20,44]. Current health status, being chronically ill and subjective life expectancy, was not associated with raw WTP or WTP per QALY (regressions not presented here). The R^2 were low, similar to related work [21].

Discussion

Recently, more empirical research to determine the monetary value of a QALY has been called for and initiated (see e.g., [19,20,44,45]). In this context, we estimated the first monetary value of a QALY in The Netherlands, using a comprehensive valuation exercise from the individual perspective. The results show that the maximum WTP for a QALY, derived through aggregating and averaging individual responses, is €24,500.

As we have shown, however, the estimates of the WTP per QALY can vary substantially, depending on the specific subgroups and methods of calculation. In terms of the latter, using the VAS valuations of the health changes rather than the TTO tariffs, resulted in an average estimate of €12,900. Such a discrepancy between TTO-based values and VAS-based values of the maximum WTP per QALY has been noted before [21,46]. Indeed, the two techniques are known to yield different estimates, but the debate regarding their acceptability or accuracy is well beyond the scope of this article (see e.g., [47] for a further discussion on VAS). For the current purpose, we consider the EuroQoL-5D tariffs to be more relevant because they are most commonly used and derived in a standardized way. In terms of the former, valuations of subgroups stratified by income level and level of certainty, proved to differ substantially. The richest, most certain subgroup elicited a considerably higher WTP per QALY (i.e., €75,400; their upper-bound estimate using the PS is €114,900). It seems important to stress these variations and to be careful with terms as "the value" of a QALY. Similarly, the SDs around our WTP estimates were considerable, indicating a large variation in preferences. The level of variation is lowest for the lower-bound estimates of WTP (i.e., highest amount people cer-

tainly would pay, as indicated on the PS) and increases with the size of WTP. Moreover, it is also important to note that the individual valuations can be combined in different ways to come to a value of a QALY. Aggregating in a different way than the one chosen here (i.e., taking the mean of ratios) is likely to result in different estimates of WTP per QALY. Such methodological aspects of deriving monetary values from the "raw material" deserve more attention, especially because there appears to be no guidance or consensus on this topic.

The results presented here align with the relevant range of the cost-effectiveness threshold of £20,000 to £30,000 (or €23,300 to €35,000) used by NICE in recent years [2] and the most commonly cited threshold of €20,000 in The Netherlands (e.g., [48]). Similar results were recently derived from the existing value of preventing a statistical fatality in the UK context, with estimates ranging between £23,199 (€26,877) and £40,029 (€46,375) per QALY [44]. Gyrd-Hansen ([20]), using a DCE approach and TTO utilities, estimated a WTP per QALY of €12,000 in the general population of Denmark for relatively small-sized health gains. King et al. [21] reported on WTP per QALY ratios obtained in three distinct patient populations. Using VAS, Standard Gamble, and TTO to elicit utilities, they found a maximum WTP per QALY ranging from $12,500 (€9500) to $32,000 (€24,500). Recently, Shiroiwa et al. [4] estimated the WTP for an additional year of survival in full health, and found that the mean WTP per QALY ranged from £23,000 (€26,600) in the UK, AU$64,000 (Australia; €36,600) to US$62,000 (US; €44,000). Seemingly, the available empirical estimates range roughly between €10,000 and €45,000—aligning with the lower- and upper-bound estimates for the full sample obtained in the current study.

We note that our results may have been influenced by several methodological issues that deserve attention. First is the range of values offered on the PS [49,50]. We carefully pretested the range of the scale in a pilot study and, to minimize the mid-point bias, employed a two-question procedure in using the PS. The majority of values chosen on the scale fell between €50 and €200. The end-point bias could be rejected, because only a few respondents (in less than 1.5% of scenarios) opted for the highest amount offered on the scale (i.e., €2500). Nevertheless, some concerns remain with the range of the scale because the results could not

be compared to results from a scale of a different range. The range may thus be limiting both the WTP estimates and the difference between WTP per QALY values stemming from the PS and the "bounded" OE question. Nevertheless, the use of a two-step procedure in deriving WTP estimates proved feasible and helpful, yet it must be noted that using other elicitation techniques may result in different estimates of WTP. This seems an important area for further research.

Furthermore, it could be argued that the respondents' ability to pay constrained the monetary value of a QALY, especially in light of nonmarginal health gains employed. Nevertheless, the data show that this too could only be a minor problem. The average maximum WTP for an average gain of 0.33 (VAS) or 0.32 (EQ-5D tariffs) was €174 a month, while the average household income was €2564. (The figure €24,500 was calculated based on the formula WTP per QALY * 12 for each individual scenario. Taking the average WTP of €174 and multiplying it by 12/0.32 would lead to a more conservative estimate that does not take the distribution of individual valuations into account). This suggests that the respondents were not bidding to the point of a catastrophic payment (i.e., on average spending 6.79% of monthly income) or that the ability to pay limited the expressed WTP. Still, employing marginal gains would decrease such concerns even further, which could result in higher estimates of WTP. In that sense, our estimates might be seen as lower bounds of the WTP per QALY [1]. Nevertheless, using marginal increments could raise questions regarding the extrapolation of obtained WTP estimates to scenarios in which nonmarginal health improvements are relevant [20]. Second, the respondents only valued potential health improvements and not potential mortality reduction. While many health-care interventions are indeed aimed at improving quality of life rather than reducing mortality, which emphasizes the relevance of the here presented figures, obtaining estimates in the context of mortality reduction remains important. It is likely that such estimates would be higher than the ones presented here [16,44]. Such scenarios would, unlike the ones used here, require valuations under risk rather than under certainty. While this has advantages such as marginality when using small risks, it also entails disturbing elements like risk weighting, which would need to be accounted for in analyzing the results. Third, the choice of the payment vehicle and frequency could be another important contextual determinant of the size of WTP per QALY estimate. In this study, WTP was elicited in relation to actual use of the intervention (as opposed to an "insurance" context) and payments phrased as monthly outlays (as opposed to, e.g., a lump sum). Although individuals in The Netherlands are somewhat acquainted with paying out-of-pocket for health care (according to OECD [51], 8% of all health care is financed out of pocket) and increasingly so since the introduction of a mandatory deductible [52], it remains unclear to what extent the type of payment seemed realistic to respondents. A similar issue has been reported by [53], and it limits the applicability of the results. Phrasing the payments in terms of a lump sum might have produced more conservative estimates because it does not offer the opportunity of spreading the burden over time [54], but could induce problems of ability to pay and budget constraints. Fourth, the relative position of the respondent's own health and the health states valued could have affected the WTP valuations because some health states could have been considered as a relative gain or a loss. Although most respondents' own health was evaluated higher than the health states presented in the questionnaire, in the analysis, we took the usual assumption that, given the instructions, own health is, in fact, irrelevant for the valuation. Our data did not allow a firm test

of whether this assumption actually holds. Similarly, the style or framing of the WTP question (i.e., valuing gains in health as opposed to valuing avoiding a loss in health) could have had an effect on WTP. Although these issues, and the scale of their impact, are beyond the scope of this study, they remain interesting empirical and theoretical questions. Finally, an important limitation of this study (and other preference elicitation studies) is the hypothetical nature of the exercise. Similar to other studies, the respondents might have found it difficult to imagine being in a health state other than what they have experienced. The same holds for other elements of our questionnaire (i.e., the concept of painless cure or the duration of health loss of precisely 12 months). Regardless of the effort put into increasing realism and reducing the hypothetical bias (i.e., through ex ante and ex post-mitigation), it is uncertain whether the elicited WTP corresponds to the real (i.e., revealed) WTP. Some have indicated that the subgroup of most certain respondents would produce an estimate of WTP that is fairly close to "real" WTP [55]. In this study, this implies that a slightly higher estimate of WTP per QALY—€26.800 instead of €24.500—would be relevant (Table 2). Given that the two estimates are not considerably different, that only one is informed by all respondents and that it has been recommended [35] to use conservative estimates in contingent valuation studies, we have focused here on the estimate of €24.500. Encouragingly, as seen above, our results compare well to the relevant range of the most often cited cost-effectiveness threshold.

Some issues need to be addressed concerning the use of the figures presented here in the context of health policy. First, the figures are lower than some thresholds that have been mentioned elsewhere, such as the $50,000 or $100,000 threshold in the US context or, for instance, the upper limit of €80,000 proposed in The Netherlands by an important advisory body RVZ [56]. Moreover, they are lower than what we infer from a part of the value-of-life literature (e.g., [57]). Our estimates are also lower than the estimates of WTP for a life-year saved [58]. Such large variations, undoubtedly fuelled by underlying methodological differences, exist and need to be explicitly addressed before recommending the use of particular thresholds in health policy. One of the key normative and methodological issues is the perspective from which the appropriate height of the cost-effectiveness threshold needs to be determined [13,59]. This appears to be an essential element in future theoretical and empirical work. For instance, the €80,000 threshold in The Netherlands was not proposed as a fixed threshold, but as the maximum of a range. Importantly, this range does not increase with individual valuations (such as income in our study), yet is increasing with more socially driven considerations, namely disease severity, which can be seen as an equity consideration [60]. NICE, perhaps somewhat similar, asks that a technology with the ICER of over £20,000 per QALY needs to make explicit references to "the particular features of the condition and population receiving the technology" as to increase its chances of being accepted [61]. Recently, NICE even indicated that certain interventions (i.e., lifesaving cancer drugs) may be approved in spite of less favorable cost-effectiveness [62,63].

Equity considerations thus appear to play a role in societal decisions. Indeed, when looking at the literature regarding "equity weights," this becomes even clearer (e.g., [17,64]). People attach weights to health gains according to the "particular features of the condition and population receiving the technology," it seems. Importantly, though, such preferences are most likely not reflected in individual valuations of own health gains. This raises the issue of usefulness of individual valuations in the current context.

If we are, however, to consider individual valuations of health gains as relevant for societal decisions on the allocation of health-care resources, the question of how to use these individual valuations is important. For example, in this study, we find a great variation in WTP across income groups. The average individual threshold of €24,500 is therefore only the "right" valuation for a small group of individuals. From a traditional normative welfare economic viewpoint, it is easy to argue that simply taking the average can result in systematically "wrong" decisions. Indeed, applying such a cost-effectiveness threshold "unduly" restricts the provision of expensive interventions to the rich part of the population (as true benefits are higher than projected), while it "unduly" grants them to poorer groups (because the true benefits are lower than projected). A similar argument extends to the use of the mean monetary value of premature fatality (i.e., value of a statistical life) as a threshold in economic evaluations done by the UK Department of Transport.

This practice resembles an implicit weighting procedure (of valuations), from an individual perspective. While we may wish to do so for many reasons, welfarist or extra-welfarist [65], the justification of using the average (as opposed to the median or the maximum) needs to be clear, and still is not (e.g., [66]).

If, on the contrary, we expect that individual valuations of own health gains may not be directly relevant for the societal decisions we are faced with, it may be worthwhile attempting to directly elicit something like the "societal WTP for a QALY." Such an ex-ante value should be the focus of future research. In our view, it should include aspects like option value and solidarity, and would be allowed to vary with characteristics of the beneficiaries of health-care interventions such as disease severity and age (instead of income). Indeed, in the context of the collective decisions in the health-care sector involving (risk and income) solidarity and other-regarding preferences, one may consider valuations directly derived from a societal perspective to be more relevant for the question at hand. This then allows a direct link between equity weights and the value of QALY gains, as well as a transparent public discourse (if not consensus) on what the desirable weights should be. It does require, however, that such valuation studies are appropriately designed, also in order to be able to interpret the results straightforwardly.

It seems, therefore, that the quest of finding appropriate monetary values is just beginning. While this study hopes to have contributed in this quest, it is clear that important normative and methodological issues need to be addressed before the results can be used in a policy context.

References

1 Johannesson M, Meltzer D. Editorial: some reflections on cost-effectiveness analysis. Health Econ 1998;7:1–7.

2 McCabe C, Claxton K, Culyer A. The NICE cost-effectiveness threshold: what it is and what that means. Pharmacoeconomics 2008;26:733–44.

3 Devlin N, Parkin D. Does NICE have a cost effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Econ 2004;13:437–52.

4 Shiroiwa T, Sung Y, Fukuda T, et al. International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness? Health Econ 2009;19:422–37.

5 Appleby J, Devlin N, Parkin D. NICE's cost effectiveness threshold. BMJ 2007;335:358–9.

6 Culyer AJ. Introduction. In: Towse A, Pritchard C, Devlin N, eds. Cost-Effectiveness Threshold: Economic and Ethical Issues. London: King's Fund and Office of Health Econ, 2002: 9–15.

7 Culyer AJ, McCabe CJ, Briggs AH, et al. Searching for a threshold not setting one: the role of the National Institute for Health and Clinical Excellence. J Health Serv Res Policy 2007;12:56–8.

8 Rawlins MD, Culyer AJ. National Institute for Clinical Excellence and its value judgments. BMJ 2004;329:224–7.

9 Eichler HG, Kong SX, Gerth WC, et al. Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge? Value Health 2004;7:518–28.

10 Dolan P, Edlin R. Is it really possible to build a bridge between cost–benefit analysis and cost-effectiveness analysis? J Health Econ 2002;21:827–43.

11 Hammitt JK. QALY versus WTP. Risk Anal 2002;22:985–1001.

12 Gyrd-Hansen D. Willingness to pay for a QALY: theoretical and methodological issues. Pharmacoeconomics 2005;23:423–32.

13 Smith RD, Richardson J. Can we estimate the "social" value of a QALY? Four core issues to resolve. Health Policy 2005;74:77–84.

14 Van Houtven G, Powers J, Jessup A, Yang J. Valuing avoided morbidity using meta-regression analysis: what can health status measures and QALYs tell us about WTP? Health Econ 2006;15:775–95.

15 Pinto-Prades JL, Loomes G, Brey R. Trying to estimate a monetary value for the QALY. J Health Econ 2009;28:553–62.

16 Bleichrodt H, Quiggin J. Life-cycle preferences over consumption and health: when is cost-effectiveness analysis equivalent to cost–benefit analysis? J Health Econ 1999;18:681–708.

17 Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. Health Econ 2005;14:197–208.

18 O'Brien B, Gafni A. When do the "dollars" make sense? Toward a conceptual framework for contingent valuation studies in health care. Med Decis Making 1996;16:288–99.

19 Weinstein MC. How much are Americans willing to pay for a quality-adjusted life year? Med Care 2008;46:343–5.

20 Gyrd-Hansen D. Willingness to pay for a QALY. Health Econ 2003;12:1049–60.

21 King JT Jr, Tsevat J, Lave JR, Roberts MS. Willingness to pay for a quality-adjusted life year: implications for societal health care resource allocation. Med Decis Making 2005;25:667–77.

22 Byrne MM, O'Malley K, Suarez-Almazor M. Willingness to pay per quality-adjusted life year in a study of knee osteoarthritis. Med Decis Making 2005;25:655–66.

23 Jansen SJT, Stiggelbout AM, Wakker PP, et al. Unstable preferences: a shift in valuation or an effect of the elicitation procedure? Med Decis Making 2000;20:62–71.

24 Smith RD. The relative sensitivity of willingness-to-pay and time-trade-off to changes in health status: an empirical investigation. Health Econ 2001;10:487–97.

25 Donaldson C, Shackley P, Abdalla M, Miedzybrodzka Z. Willingness to pay for antenatal carrier screening for cystic fibrosis. Health Econ 1995;4:439–52.

26 Donaldson C, Shackley P, Abdalla M. Using willingness to pay to value close substitutes: carrier screening for cystic fibrosis revisited. Health Econ 1997;6:145–59.

27 Donaldson C, Thomas R, Torgerson D. Validity of open-ended and payment scale approaches to eliciting willingness to pay. App Econ 1997;29:79–84.

28 Olsen JA, Donaldson C. Helicopters, hearts and hips: using willingness to pay to set priorities for public sector health care programmes. Soc Sci Med 1998;46:1–12.

29 Gibb S, Donaldson C, Henshaw R. Assessing strength of preference for abortion method using willingness to pay. J Adv Nurs 1998;27:30–6.

30 Dubourg WR, Jones-Lee MW, Loomes G. Imprecise preferences and survey design in contingent valuation. Economica 1997; 64:681–702.

31 Bhatia MR, Fox-Rushby JA. Willingness to pay for treated mosquito nets in Surat, India: the design and descriptive analysis of a household survey. Health Policy Plan 2002;17:402–11.

32 Bhatia MR, Fox-Rushby JA. Validity of willingness to pay: hypothetical versus actual payment. Appl Econ Lett 2003;10:737–40.

33 Johnson R, Banzhaf MR, Desvousges WH. Willingness to pay for improved respiratory and cardiovascular health: a multiple-format, stated-preference approach. Health Econ 2000;9: 295–317.

34 Blomquist G, Blumenschein K, Johannesson M. Eliciting willingness to pay without bias using follow-up certainty statements: comparisons between probably/definitely and a 10-point certainty scale. Environ Resour Econ 2009;43:473–502.

35 Arrow K, Solow R, Portney P, et al. Report of the NOAA Panel on contingent valuation. Fed Reg 1993;58:4602–14.

36 Smith RD. It's not just what you do, it's the way that you do it: the effect of different payment card formats and survey administration on willingness to pay for health gain. Health Econ 2006;15:281–93.

37 Blumenschein K, Johannesson M, Yokoyama KK, Freeman PR. Hypothetical versus real willingness to pay in the health care sector: results from a field experiment. J Health Econ 2001;20:441–57.

38 Poe G, Clark JE, Rondeau D, Schulze WD. Provision point mechanisms and field validity tests of contingent valuation. Environ Resour Econ 2002;23:105–31.

39 Bateman IJ, Jones AP. Contrasting conventional with multi-level modeling approaches to meta-analysis: expectation consistency in U.K. woodland recreation values. Land Econ 2003;79:235–58.

40 Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. BMJ 1998;316:736–41.

41 Lamers LM, McDonnell J, Stalmeier PF. The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. Health Econ 2006;15:1121–32.

42 CBS. Available from: http://www.cbs.nl/en-GB/menu/cijfers/kerncijfers/default.htm [Accessed March 21, 2009].

43 Zethraeus N. Willingness to pay for hormone replacement therapy. Health Econ 1998;7:31–8.

44 Mason H, Jones-Lee M, Donaldson C. Modelling the monetary value of a QALY: a new approach based on UK data. Health Econ 2008;18:933–50.

45 EUROVAQ project. Available from: http://research.ncl.ac.uk/eurovaq/ [Accessed May 27, 2009].

46 Lundberg L, Johannesson M, Silverdahl M, et al. Quality of life, health-state utilities and willingness to pay in patients with psoriasis. Br J Dermatol 1999;141:1067–75.

47 Parkin D, Devlin N. Is there a case for using visual analogue scale valuations in cost–utility analysis? Health Econ 2006;15:653–64.

48 Brouwer WBF, van Exel NJA, Baker R, Donaldson C. The new myth: the social value of the QALY. Pharmacoeconomics 2008; 26:1–4.

49 Whynes DK, Wolstenholme JL, Frew EJ. Evidence of range bias in contingent valuation payment scales. Health Econ 2004;13:183–90.

50 Frew EJ, Whynes DK, Wolstenholme J. Eliciting willingness to pay: comparing closed-ended with open-ended and payment scale formats. Med Decis Mak 2003;23:150–9.

51 OECD statistics. Available from: http://stats.oecd.org/WBOS/index.aspx [Accessed April 2, 2009].

52 Holland J, van Exel NJA, Schut FT, Brouwer WBF. Some pain, no gain. Experiences with the no-claim rebate in the Dutch health care system. Health Econ Policy Law 2009;4:405–24.

53 Skjoldborg US, Gyrd-Hansen D. Conjoint analysis. The cost variable: an Achilles' heel? Health Econ 2003;12:479–91.

54 Carson RT. Contingent valuation: a user's guide. Environ Sci Technol 2000;34:1413–8.

55 Blumenschein K, Blomquist GC, Johannesson M, et al. Eliciting willingness to pay without bias: evidence from a field experiment. Econ J 2008;118:114–37.

56 RVZ. Available from: http://www.rvz.net/cgi-bin/rvz_p.pl?id=54 [Accessed June 12, 2009].

57 Hirth RA, Chernew ME, Miller E, et al. Willingness to pay for a quality-adjusted life year: in search of a standard. Med Decis Making 2000;20:332–42.

58 Braithwaite RS, Meltzer DO, King JT Jr, et al. What does the value of modern medicine say about the $50,000 per quality-adjusted life-year decision rule? Med Care 2008;4:349–59.

59 Dolan P, Olsen JA, Menzel P, Richardson J. An inquiry into the different perspectives that can be used when eliciting preferences in making. Health Econ 2003;12:545–51.

60 Stolk EA, van Donselaar G, Brouwer WBF, van Busschbach JJ. Reconciliation of economic concerns and health policy: illustration of an equity adjustment procedure using proportional short-fall. Pharmacoeconomics 2004;22:1097–107.

61 NICE. Guide to the Methods of Technology Appraisal. 2004 (reference N0515, Available from: http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyapprais alprocessguides/guide_to_the_methods_of_technology_ appraisal_reference_n0515.jsp [Accessed April 27, 2009].

62 NICE. Appraising life extending, end of life treatments. 2008. http://www.nice.org.uk/media/88A/F2/ SupplementaryAdviceTACEoL.pdf [Accessed July 20, 2009].

63 Towse A. Should NICE's threshold range for cost per QALY be raised? Yes. BMJ 2009;338:181.

64 Williams A, Cookson R. Equity in Health. Handbook of Health Economics. Culyer AJ, Newhouse JP, eds. North Holland, Amsterdam: Handbook of Health Economics, 2000.

65 Brouwer WBF, Culyer AJ, van Exel NJA, Rutten FFH. Welfarism vs. extra-welfarism. J Health Econ 2008;2:325–38.

66 Baker R, Chilton S, Jones-Lee M, Metcalf H. Valuing lives equally: defensible premise or unwarranted compromise? J Risk Uncertain 2008;36:125–38.