Willingness to Pay for Health-Related Quality of Life Gains in Relation to Disease Severity and the Age of Patients

Vivian Reckers-Droog, MSc, Job van Exel, PhD, Werner Brouwer, PhD

A B S T R A C T

Objectives: Decision-making frameworks that draw on economic evaluations increasingly use equity weights to facilitate a more equitable and fair allocation of healthcare resources. These weights can be attached to health gains or reflected in the monetary threshold against which the incremental cost-effectiveness ratios of (new) health technologies are evaluated. Currently applied weights are based on different definitions of disease severity and do not account for age-related preferences in society. However, age has been shown to be an important equity-relevant characteristic. This study examines the willingness to pay (WTP) for health-related quality of life (QOL) gains in relation to the disease severity and age of patients, and the outcome of the disease.

Methods: We obtained WTP estimates by applying contingent-valuation tasks in a representative sample of the public in The Netherlands (n = 2023). We applied random-effects generalized least squares regression models to estimate the effect of patients' disease severity and age, size of QOL gains, disease outcome (full recovery/death 1 year after falling ill), and respondent characteristics on the WTP.

Results: Respondents' WTP was higher for more severely ill and younger patients and for larger-sized QOL gains, but lower for patients who died. However, the relations were nonlinear and context dependent. Respondents with a lower age, who were male, had a higher household income, and a higher QOL stated a higher WTP for QOL gains.

Conclusions: Our results suggest that—if the aim is to align resource-allocation decisions in healthcare with societal preferences—currently applied equity weights do not suffice.

Keywords: age, end of life, health-related quality of life, severity of illness, willingness to pay.

VALUE HEALTH. 2021; ■(■):■–■

Introduction

An important objective of publicly financed healthcare systems is to maximize population health given a certain budget constraint.[1] To meet this objective, economic evaluations can be used to inform decision makers about whether reimbursing a (new) health technology can be considered good value for money. In economic evaluations, health gains are often expressed in terms of quality-adjusted life-years (QALYs), comprising gains in both health-related quality of life (QOL) and life expectancy (LE).[2,3] The incremental cost-effectiveness ratio (ICER) of a technology is evaluated against a monetary threshold that represents the maximum societal willingness to pay (WTP) for a QALY or the opportunity costs of spending within the healthcare sector.[4-6]

Traditionally, a "QALY is a QALY is a QALY" in economic evaluations,[7] meaning that all health gains are valued equally. However, equity weights can be attached to health gains or reflected in the monetary threshold to facilitate a more equitable and fair allocation of healthcare resources.[1,8-12] In the former case, the

equity-adjusted ICER of a technology is evaluated against a fixed monetary threshold and in the latter case, the (unadjusted) ICER of a technology is evaluated against a flexible, equity-adjusted monetary threshold.[1,10] These weights can be based on a range of equity considerations that, for example, are related to characteristics of the patients, disease, or technology.[1,12-20] To facilitate consistent and accountable decision making, it has been advocated to explicitly and transparently integrate such considerations into the decision-making framework.[21-24] Although many countries (eg, France, Germany, Sweden, and Australia) do this in an ad hoc, implicit manner,[25-27] Norway, The Netherlands, and England do this in an explicit manner by applying equity weights.[28-31] Text Box 1 includes a brief overview of how the weights are applied in these countries.

Societal preferences for equity weighting based on disease severity (defined broadly here to include absolute shortfall, proportional shortfall, and end-of-life considerations associated with terminal illnesses as described in Text Box 1) are increasingly studied, also in relation to patients' age. The available evidence

---

TEXT BOX 1. Application of equity weights in Norway, The Netherlands, and England.

In Norway and The Netherlands, the (unadjusted) incremental cost-effectiveness ratio (ICER) of a health technology is evaluated against a monetary threshold that is weighted based on the disease severity of the targeted patient population.[11,28,31] In Norway, a flexible threshold in the range of NOK 275 000 to 825 000 (~€27 500 to €82 500) per quality-adjusted life-year (QALY) is (informally) applied, with a maximum weight of 3 attached for evaluating the ICER of a health technology that targets patients with the highest level of disease severity (ie, an absolute shortfall of ≥20 QALYs).[31–33] Absolute shortfall is calculated as the disease-related loss of remaining QALYs without the new health technology, compared to the remaining QALY expectation in absence of the disease.[31,34,35] In The Netherlands, a flexible threshold in the range of €20 000 to €80 000 per QALY gained is applied, with a maximum weight of 4 attached for evaluating the ICER of a health technology that targets patients with the highest level of disease severity (ie, a proportional shortfall of 0.71–1.00).[11,28,29] Proportional shortfall is calculated as the proportion of absolute shortfall, relative to the remaining QALY expectation in absence of the disease and measured on a scale from 0 "no QALY loss" to 1 "complete QALY loss".[8,28] Health technologies that target patients with the lowest level of disease severity (ie, a proportional shortfall of <0.10) are generally not recommended for reimbursement.[28,36,37] In England, a weight in the range of 1.7 to 2.5 can be attached to QALYs that are gained by prolonging the lives of terminally ill patients (normally with a remaining life expectancy of ≤24 months) by at least 3 months.[30] The resulting (equity-adjusted) ICER of a health technology is then evaluated against the common threshold range of £20 000 to £30 000 (~€23 000 to €34 500).[30] Note that the weights are currently applied in decisions on the reimbursement of new health technologies, not in decisions on the displacement of other technologies that may follow as a result (eg, in England[30,38]).

---

suggests that the public considers age to be an important equity-relevant characteristic (often reflected by giving a higher weight to health gains in younger patients), possibly even more important than disease severity.[1,5,11,39–42] Nevertheless, the weights applied in Norway, The Netherlands, and—classifying terminally ill patients as severely ill[43]—England are all based on disease severity. These weights do not directly account for patients' age nor aim to weight age in resource-allocation decisions, even though they may be inextricably related to patients' age.[11] For example, the weights based on absolute shortfall in Norway may implicitly prioritize younger over older patients, and, conversely, the weights based on proportional shortfall in The Netherlands and end-of-life considerations in England may implicitly prioritize older over younger patients.[11,44–46]

The aim of this study was to examine the willingness to pay (WTP) for health gains in relation to the disease severity and age of patients and to examine whether the WTP was different between health gains in patients who fully recovered and patients who died (1 year after falling ill). Based on the available evidence on societal preferences in this context, we hypothesized that WTP would be higher for more severely ill and younger patients, and for patients who can be considered terminally ill. We further hypothesized that the elicited WTP would be sensitive to scale and to household income, indicating the theoretical validity of the elicited WTP.[47]

We elicited the WTP for health gains in terms of an increase in monthly basic health-insurance premium in a representative sample of the general public in The Netherlands, as this relates directly to the payment vehicle used for collectively funding healthcare in this country. Given the aim of our study, we focus on the relative rather than absolute WTP for health gains. The results of this study may inform decisions on the relative size of severity-and/or age-dependent equity weights and on the range and shape of monetary thresholds used to evaluate the ICERs of health technologies. The results are considered to be of particular interest to Norway, The Netherlands, and England given their current use of equity weights, but also to other countries that (intend to) integrate equity and efficiency considerations into their formal decision-making framework.

Methods

Sample and Data Collection

We designed a contingent-valuation (CV) study that was administered online by a professional research agency (Dynata). Respondents were quota sampled to be representative of the general public in The Netherlands in terms of age (18-75 years), sex, and education level and to obtain a broad range in household income. Before conducting the main study in August 2019, we conducted a pilot study in a small sample (n = 100) to test the range of the payment scale and clarity of the tasks. The results of this study did not lead to modifications, and hence we merged the pilot and main data before conducting the analyses (total sample n = 2023).

Before respondents completed the questionnaire, we explained that healthcare resources are scarce and decision makers use information on societal preferences to allocate the available resources in an optimal manner for society. We asked respondents to complete the CV tasks from a socially-inclusive-personal (SIP) perspective.[48] This implied that they had to take into consideration the possibility that they themselves, their family, friends, and/or acquaintances could be part of the hypothetical patient group as well as unknown others. As the SIP perspective represents a combination of the personal and social perspectives,[48] applying it may be seen as yielding relevant WTP estimates for health gains in the context of a collectively funded healthcare system like that of The Netherlands.[48,49] Upon completion of the questionnaire, respondents received a fee of 50 eurocents that they could save in a personal account or donate to charity.

Questionnaire

The questionnaire consisted of 4 parts. In part 1, we introduced respondents to the following concepts using text and graphs: (1) QOL, operationalized on a visual analogue scale (VAS) ranging from 0 "dead" to 100 "full health"; (2) disease severity, operationalized as disease-related QOL loss (in points from 100 on the VAS) in patients who fully recovered and as a combination of QOL and LE (in years) loss in patients who died 1 year after falling ill;

TEXT BOX 2. Contingent-valuation procedure.

Willingness to pay was elicited by applying a 2-step contingent-valuation procedure, consisting of a payment scale and a bounded open-ended question.[50] The payment scale ranged from a €0 to €24 increase in monthly basic health-insurance premium with unevenly distributed intervals between the value points (ie, €0, €0.50, €1, €1.50, €2, €2.50, €3, €4, €5, €6, €7, €8, €10, €12, €14, €16, €18, €20, €22, €24, and "more"). Note that monthly payment of health-insurance premiums is mandatory for adults (18+) in The Netherlands. By approximation, the number of adults was 13.7 million and the monthly basic health-insurance premium was €115.00 per person in 2019.[51] In step 1, we asked respondents to inspect the payment scale from left to right and indicate the increase in monthly premium they were certainly willing to pay for the duration of 1 year. We then asked them to again inspect the payment scale from left to right and indicate the increase in monthly premium they would certainly *not* be willing to pay for the duration of 1 year. In step 2, we asked respondents to indicate the maximum increase in monthly premium they would be willing to pay within the range obtained in step 1. In both steps, we asked respondents to take their net monthly household income into account as a proxy for their ability to pay. We asked respondents who stated a WTP of €0 in step 1 to explain their main reason for having this preference by completing an open-text field or checking 1 of 6 randomized answer options. Three answer options related to true zero valuations (ie, "I cannot afford to pay more than €0", "Treating these patients is not worth more than €0 to me", and "I believe the treatment is worth more than €0, but I would rather spend my money on something else"), and 3 answer options related to protest zero valuations (ie, "I am against an increase in monthly basic health-insurance premium", "Patients should pay for the treatment themselves", and "The value of health and healthcare cannot be expressed in monetary terms"). The open-text field answers were qualitatively assessed by the first 2 authors and subsequently classified as either a true or protest zero valuation.

and (3) treatment-related QOL gain (in points on the VAS). We familiarized respondents with the concepts and tasks by asking them to assess their own QOL "today" on the VAS and complete a practice task from a personal perspective.[48] After completing this task, we asked respondents to assess its level of clarity on a 7-point Likert scale ranging from 1 "very unclear" to 7 "very clear" and to indicate on what they would likely economize to cover the stated WTP to increase their awareness of the associated opportunity costs.

In part 2, respondents completed 2 tasks from a SIP perspective for which they were randomly assigned to 2 out of 20 scenarios. Each scenario started with the introduction of a group of 10 000 patients aged 10, 20, 40, or 70. We explained that the patients would have lived in full health (a score of 100 on the VAS) until the age of 80 had they not fallen ill. Due to the disease, their QOL decreased from 100 to either 90, 70, 50, 30, or 10 on the VAS for the duration of 1 year. After this year, they would fully recover (ie, a score of 100 on the VAS). The disease would not affect their LE. We explained that a treatment was available that would increase patients' QOL with 10 points on the VAS during the year of illness and that the treatment type and costs were the same for all patients. The treatment could be made available to patients by increasing the monthly basic health insurance premium for the duration of 1 year. This increase would apply to all adult inhabitants of The Netherlands. We elicited respondents' WTP for the treatment-related QOL gains by applying the 2-step procedure described in Text Box 2. After respondents completed the second task, we asked them to indicate how certain they were of actually paying the stated WTP in case the increase became effective immediately, on a 7-point Likert scale ranging from 1 "very uncertain" to 7 "very certain".

In part 3, respondents completed 1 additional task from a SIP perspective for which they were randomly assigned to 1 out of 20 scenarios that were evenly distributed across 5 modules. In modules 1–4, the level of disease severity was 50, treatment-related QOL gain was 20 points (modules 1 and 3) or 50 points (modules 2 and 4), and patients either fully recovered (modules 1 and 2) or died (modules 3 and 4) 1 year after falling ill. We used the data from modules 1 and 2, and from the scenarios in part 2 in which the level of disease severity was also 50 to examine whether respondents' WTP was sensitive to scale.[47] We used the

data from modules 1–4 to examine whether respondents' WTP for similar-sized QOL gains was different between gains in patients who fully recovered and patients who died. In module 5, we focused on a different question that is reported elsewhere. Appendix A includes a task example and Appendix B an overview of the scenario characteristics (see Supplemental Materials found at https://doi.org/10.1016/j.jval.2021.01.012).

In part 4, we asked respondents about their sociodemographic characteristics.

Statistical Analyses and Hypotheses

Before conducting the analyses, we excluded protest zero valuations (see Text Box 2), outliers, and speeders. We identified outliers based on the distribution of stated WTPs (z-score ≥1.64). We classified respondents who completed the 3 tasks in less than 90 seconds as speeders, based on a timed test of completing the tasks by 3 researchers not involved in this study.

We calculated the raw mean (SD; 95% CI) WTP for QOL gains in all scenarios and the difference in raw mean (SE; 95% CI) WTP for similar sized QOL gains in patients who fully recovered and patients who died 1 year after falling ill. We applied 2-tailed Welch's t-tests (Bonferroni corrected) to examine whether the latter was statistically significantly different from 0. Furthermore, we applied 7 random-effects generalized least squares models to estimate the effect of scenario and respondent characteristics on the WTP. Models 1 and 2 were based on the data obtained in part 2 of the questionnaire and included the scenario characteristics disease severity and age of patients, and their interaction. Models 3 to 6 were based on the data obtained in part 2 *and* 3 of the questionnaire. Model 3 included the scenario characteristics disease severity, age of patients, size of QOL gains, and disease outcome (full recovery/death, 1 year after falling ill). Model 4 also included the interaction between the disease severity and age of patients. Models 5 to 7 consecutively included the interactions between the disease severity and age of patients, size of QOL gains and age of patients, and disease outcome and age of patients as well as the respondent characteristics age, age^2, sex, children (yes/no), education level, household income (adjusted for household size using an elasticity scale of 0.5 to account for economies of scale[52]), and QOL.

We assumed that respondents might have used their first WTP stated from a SIP perspective as a reference (anchor) in the subsequent tasks. After testing this assumption, we decided to account for a time effect in all models (labeled "CV task").[53] A downward adjustment of a previously stated WTP could also indicate a violation of the monotonicity principle that a larger-sized QOL gain should, ceteris paribus, result in a higher WTP.[54] Therefore, we performed sensitivity analyses to examine the robustness of our results by repeating the analyses excluding respondents who, ceteris paribus, stated a lower WTP for larger QOL gains. We also examined the robustness of our results by alternately repeating the analyses excluding respondents who reported a low clarity score (ie, 1−3 level) for the practice task, reported a low certainty score (ie, 1−3 level) for actually paying the stated WTP, and completed the 3 tasks in less than 39 (instead of the predetermined 90) seconds based on the distribution of completion times (z-score ≤-1.64). Furthermore, we examined the effect of respondents' proximity to the age of patients and of respondents' stated WTP in the practice task on the WTP.

Before conducting the analyses, we hypothesized that respondents' WTP would be higher for QOL gains in more severely ill patients (ie, patients with a higher level of disease severity and patients who died 1 year after falling ill) and for QOL gains in younger patients. Moreover, we hypothesized that respondents' WTP would be sensitive to scale and to household income in the sense that the WTP would be higher for larger-sized QOL gains and for respondents with a relatively higher household income. Evidence in support of the latter hypothesis would indicate the theoretical validity of the elicited WTP.[47]

We conducted the analyses using Stata 16.1 (Stata Corp LP, College Station, TX).

Results

Table 1 presents the descriptive statistics of the sample (n = 1317) that remained after excluding protest zero valuations (n = 73), outliers (n = 31), and speeders (n = 602). Of the speeders, 50 also gave protest zero valuations, and 12 also stated an outlying WTP. The statistics indicate that the sample was representative of the general public in The Netherlands in terms of sex and education level, but somewhat older.

The remaining respondents assessed the mean (SD) clarity of the practice task at 5.9 (1.1) and certainty of actually paying the stated WTP at 5.4 (1.3) on the 7-point Likert scale. A total of 50 (3.8%) respondents reported a low clarity score, 98 (8.3%) a low certainty score, and 37 (2.8%) stated, ceteris paribus, a lower WTP for a larger-sized QOL gain.

Table 2 presents the raw mean (SD; 95% CI) WTP for QOL gains of 10 points in patients who fully recovered 1 year after falling ill. On average, the WTP was €8.0 per month for the duration of 1 year. The results indicate that respondents' WTP was generally higher for QOL gains in more severely ill patients (average WTP: €7.3 to €8.4) and younger patients (average WTP: €7.8 to €8.4); however, the relations were nonlinear. We observed a relatively low average WTP of €8.0 for QOL gains in patients with a severity level of 50 and a relatively high average WTP of €8.4 for QOL gains in patients with a severity level of 70. We also observed a relatively low average WTP of €7.9 for QOL gains in patients aged 20. Because the SDs were relatively large and the 95% CIs largely overlapped, these results suggest strong preference heterogeneity and only partially support the hypotheses that respondents' WTP is higher for more severely ill and younger patients.

Table 1. Descriptive statistics.

	Sample (N = 1317)*		General public[†]	
	%	Mean (SD)	%	Mean
Age (Years)		51.5 (15.9)		46.1
Sex (Female)	51.3		50.3	
Education level[‡]				
Low	11.6		8.6	
Medium	58.3		57.4	
High	30.1		32.5	
Household income (After tax)				
<€1999	31.7			
€2000–€3999	42.1			
≥€4000	20.7			
NS	5.5			
Children (Yes)	60.0			
QOL (0-100 VAS)		81.3 (16.4)		
Completion time of CV tasks (minutes)		6.0 (27.7)		

CV indicates contingent valuation; NS, not stated; QOL, health-related quality of life; VAS, visual analogue scale (ranging from 0 "dead" to 100 "full health").
*Consistent protest zero valuations (n = 73), outliers (n = 31), speeders (n = 540), and respondents who met more than one exclusion criterion (n = 62) are excluded from the table.
[†]Age is based on statistics for the population aged 18-75 years, sex is based on statistics for the overall population, and education level is based on statistics for the population aged 15-75 years. Population statistics for 2019; source: Statistics Netherlands (https://opendata.cbs.nl/statline).
[‡]Low = lower vocational and primary school, Medium = middle vocational and secondary school, High = higher vocational and academic education.

Table 3 presents the raw mean (SD; 95% CI) WTP for QOL gains of 20 and 50 points in patients with a severity level of 50 who fully recovered and who died 1 year after falling ill (for comparison presented with the WTP for QOL gains of 10 points in patients with severity level 50 who fully recovered, copied from Table 2). These results indicate that respondents' WTP was generally higher for larger sized QOL gains and "hump shaped" across ages, with a peak at age 10, 20, or 40 depending on the size of the gain and whether patients fully recovered or died. Respondents' WTP was higher for similar-sized QOL gains in patients who fully recovered than in patients who died, except for gains of 20 points in patients aged 20 and 40. Most differences were not statistically significant, except for the lower mean (SE; 95% CI) WTP of €4.5 (1.8; –8.0, –1.0) for QOL gains of 50 points in patients aged 20 who died than in those who fully recovered (Bonferroni corrected, $\alpha/4$). Although these results indicate that respondents' WTP is higher for younger patients and larger-sized QOL gains, they do not support the hypothesis that respondents' WTP is higher for patients who die.

Tables 4 and 5 present the regression results. Note that models 1 and 2 are based on the data obtained in part 2, and models 3 to 7 are based on the data obtained part 2 *and* 3 of the questionnaire.

Table 2. Raw mean (SD; 95% CI) WTP for QOL gains of 10 points in patients who will fully recover 1 year after falling ill (in €).*

Age[†]	Severity[‡]										Average
	10		30		50		70		90		
	n	Mean (SD; 95% CI)	*n*	Mean (SD; 95% CI)	*n*	Mean (SD; 95% CI)	*n*	Mean (SD; 95% CI)	*n*	Mean (SD; 95% CI)	
10	123	7.8 (7.1; 6.5, 9.0)	127	8.5 (6.9; 7.3, 9.7)	134	8.2 (6.4; 7.1, 9.3)	127	9.0 (7.7; 7.7, 10.3)	121	8.4 (7.2; 7.1, 9.7)	8.4 (7.0; 7.8, 8.9)
20	118	7.0 (6.5; 5.8, 8.2)	116	7.5 (6.6; 6.3, 8.7)	121	7.3 (6.7; 6.1, 8.5)	130	9.0 (7.2; 7.7, 10.2)	112	8.7 (7.5; 7.3, 10.2)	7.9 (6.9; 7.4, 8.5)
40	118	7.0 (6.9; 5.8, 8.3)	118	8.6 (7.0; 7.4, 9.9)	129	8.5 (7.1; 7.2, 9.7)	133	8.5 (6.4; 7.4, 9.6)	120	7.7 (6.7; 6.4, 8.9)	8.1 (6.8; 7.5, 8.6)
70	116	7.5 (7.3; 6.2, 8.9)	127	8.2 (6.9; 7.0, 9.4)	128	7.8 (7.1; 6.5, 9.0)	110	7.1 (6.3; 5.9, 8.3)	125	8.2 (6.3; 7.1, 9.3)	7.8 (6.8; 7.2, 8.3)
Average		7.3 (6.9; 6.7, 8.0)		8.2 (6.8; 7.6, 8.8)		8.0 (6.8; 7.4, 8.5)		8.4 (6.9; 7.8, 9.0)		8.2 (6.9; 7.6, 8.9)	8.0 (6.9; 7.8, 8.3)

QOL indicates health-related quality of life; WTP, willingness to pay (in € per month for the duration of 1 year).
*Protest zero valuations, outliers (raw WTP ≥€35.00), and speeders are excluded from this table. In the scenarios, the groups consist of 10 000 patients who fully recover 1 year after falling ill. The treatment-related QOL gain is 10 points, measured on a visual analogue scale ranging from 0 "death" to 100 "full health".
[†]Age at onset of the disease (in years).
[‡]Severity is operationalized in terms of disease-related QOL loss and measured in points from 100 on the visual analogue scale.

The results indicate that—compared to severity level 10—a higher severity level was, ceteris paribus, associated with a higher WTP, though it was relatively low for patients with severity level 90 (model 1: β 1.08), and it decreased at a marginal rate when data were aggregated (eg, model 2: β 0.83 to 1.32). Compared to age 10, a higher age was, ceteris paribus, associated with a lower WTP (model 1: β −0.25 to β −0.84; model 3: β −0.09 to −1.09), but only the coefficient for age 70 was statistically significant. The interaction between patients' disease severity and age indicates in some scenarios that respondents' WTP for QOL gains in patients with different levels of disease severity was dependent on their age. When patients had a severity level of 10, respondents' WTP was, ceteris paribus, higher for patients aged 10 than for patients aged 20 (model 2: β -1.10; model 4: β −1.39; model 5: β −1.50) and for patients aged 40 (model 2: β −1.22; model 4: β −1.34; model 5: β −1.45). However, WTP was relatively higher for patients aged 20 from severity level 50 onward (model 2: β 1.14 to 1.63; model 4: β 1.45 to 2.23; model 5: β 1.51 to 2.39) and for patients aged 40 from severity level 70 onward (model 2: β 1.57; model 4: β 1.35 to 1.92; model 5: β 1.58 to 2.19). The results further indicate that—compared to QOL gains of 10 points—QOL gains of 20 and 50 points were, ceteris paribus, associated with a higher WTP though

Table 3. Raw mean (SD; 95% CI) WTP for QOL gains in patients who will fully recover and die 1 year after falling ill (in €).*

Age[†]	Disease outcome 1 year after falling ill										Difference	
	Full recovery						Death					
	Part 2		Part 3				Part 3				Part 3	
			Module 1		Module 2		Module 3		Module 4		Module 3 – 1	Module 4 – 2
	10 points		20 points		50 points		20 points		50 points		20 points	50 points
	n	Mean (SD; 95% CI)	*n*	Mean (SD; 95% CI)	*n*	Mean (SD; 95% CI)	*n*	Mean (SD; 95% CI)	*n*	Mean (SD; 95% CI)	ΔMean (SE; 95% CI)	ΔMean (SE; 95% CI)
10	134	8.2 (6.4; 7.1, 9.3)	66	8.5 (6.1; 7.0, 10.0)	59	10.4 (11.0; 7.6, 13.3)	61	7.5 (6.2; 5.9, 9.1)	71	9.8 (6.7; 8.2, 11.4)	-1.0 (1.1; −3.2, 1.2)	-0.6 (1.6; −3.9, 2.6)
20	121	7.3 (6.7; 6.1, 8.5)	61	8.9 (7.1; 7.1, 10.8)	66	12.3 (12.9; 9.2, 15.5)	57	9.7 (7.4; 7.7, 11.6)	64	7.8 (6.2; 6.3, 9.4)	0.7 (1.3; −1.9, 3.4)	−4.5[‡] (1.8; −8.0, −1.0)
40	129	8.5 (7.1; 7.2, 9.7)	61	7.4 (5.8; 5.9, 8.9)	61	10.1 (10.1; 7.5, 12.6)	59	7.9 (7.5; 6.0, 9.9)	61	8.9 (8.1; 6.8, 10.9)	0.6 (1.2; −1.9, 3.0)	-1.2 (1.7; −4.5, 2.1)
70	128	7.8 (7.1; 6.5, 9.0)	54	7.5 (6.4; 5.7, 9.2)	68	8.6 (9.4; 6.3, 10.8)	61	5.6 (5.1; 4.3, 6.8)	60	8.0 (6.9; 6.2, 9.8)	-1.9 (1.1; −4.1, 0.2)	-0.6 (1.4; −3.4, 2.3)

QOL indicates health-related quality of life (measured in points on a visual analogue scale, ranging from 0 "dead" to 100 "full health"); WTP, willingness to pay (in € per month for the duration of 1 year).
*Respondents with protest zero valuations, an outlying WTP (≥€35.00 in part 2, €26.50 in Module 1, ≥€63.00 in Module 2, ≥€50.00 in Module 3, and ≥€32.80 in Module 4), and who completed the tasks too quickly are excluded from this table. In the scenarios, the groups consist of 10 000 patients with a severity level of 50 points on the visual analogue scale.
[†]Age at onset of the disease (in years).
[‡]$P < .05$ (Bonferroni corrected, $\alpha = 0.05/4$).

Table 4. Regression results.[†]

DV: WTP	Model 1		Model 2		Model 3		Model 4[‡]	
	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
Severity[§] (10 = reference)	-	-	-	-	-	-	-	-
30	0.64** (0.26)	0.13, 1.16	0.41 (0.52)	−0.60, 1.42	0.83*** (0.27)	0.30, -1.37	0.40 (0.49)	−0.56, 1.36
50	1.07**** (0.27)	0.54, 1.59	0.71 (0.58)	−0.43, 1.84	1.19**** (0.27)	0.66, 1.72	0.69 (0.47)	−0.23, 1.60
70	1.23**** (0.26)	0.73, 1.73	0.44 (0.53)	−0.61, 1.48	1.22**** (0.26)	0.72, 1.73	0.54 (0.48)	−0.41 to 1.49
90	1.08**** (0.29)	0.52, 1.64	0.37 (0.55)	0.71, -1.45	1.32**** (0.29)	0.75, 1.90	0.20 (0.56)	−0.90, 1.31
Age[‖] (10 = reference)	-	-	-	-	-	-	-	-
20	−0.35 (0.23)	−0.88, 0.09	−1.11** (0.50)	−2.09, −0.12	−0.09 (0.24)	−0.56, 0.37	−1.39*** (0.52)	−2.42, −0.37
40	−0.25 (0.23)	−0.69, 0.20	−1.22** (0.54)	−2.28, −0.17	−0.24 (0.21)	−0.62, 0.19	−1.34** (0.53)	−2.38, −0.29
70	−0.84**** (0.23)	−1.30, −0.39	−0.62 (0.54)	−1.67, 0.43	−1.09**** (0.21)	−1.51, −0.68	−0.66 (0.53)	−1.70, 0.38
Severity × Age (Severity 10; Age 10 = reference)			-	-			-	-
30 × 20			0.71 (0.65)	−0.57, 1.99			1.04 (0.68)	−0.30, 2.38
30 × 40			1.06 (0.72)	−0.35, 2.47			1.20 (0.75)	−0.27, 2.66
30 × 70			−1.05 (0.75)	−2.51, 0.42			−0.64 (0.74)	−2.10, 0.81
50 × 20			0.33 (0.73)	−1.10, 1.75			1.45** (0.67)	0.13, 2.77
50 × 40			1.17 (0.79)	−0.39, 2.73			1.18 (0.65)	−0.09, 2.46
50 × 70			−0.31 (0.79)	−1.86, 1.23			−0.77 (0.64)	−2.03, 0.49
70 × 20			1.14 (0.67)	−0.19, 2.46			1.54** (0.68)	0.20, 2.87
70 × 40			1.57** (0.72)	0.16, 2.98			1.35** (0.69)	0.00, 2.70
70 × 70			0.30 (0.73)	−1.12, 1.73			−0.21 (0.69)	−1.56, 1.15
90 × 20			1.63** (0.76)	0.15, 3.12			2.23*** (0.83)	0.61, 3.86
90 × 40			1.05 (0.72)	−0.37, 2.47			1.92** (0.77)	0.41, 3.43
90 × 70			0.02 (0.76)	−1.46, 1.50			0.33 (0.77)	−1.17, 1.83
QOL gain (10 points = reference)					-	-	-	-
20 points					1.44**** (0.31)	0.82, 2.05	1.43**** (0.31)	0.82, 2.05
50 points					2.82**** (0.40)	2.04, 3.60	2.83**** (0.40)	2.04, 3.62
Death (Full recovery = reference)[¶]					−0.80** (0.34)	−1.47, −0.13	−0.79** (0.34)	−1.46, −0.12
CV task	-0.83**** (0.12)	−1.07 to −0.59	−0.81**** (0.12)	−1.05, −0.58	−0.85**** (0.13)	−1.09, −0.60	−0.84**** (0.12)	−1.08, −0.60
Constant	8.93**** (0.35)	8.25-9.62	9.32**** (0.46)	8.42, 10.23	8.89**** (0.35)	8.20, 9.57	9.39**** (0.43)	8.55, 10.24
R^2 overall	0.00		0.01		0.01		0.01	
n (observations; groups)	2521; 1317		2251; 1317		3443; 1308		3443; 1308	

CV indicates contingent valuation; DV, dependent variable; NS, not stated; QOL, health-related quality of life; VAS, visual analogue scale (ranging from 0 "dead" to 100 "full health"); WTP, willingness to pay (in € per month for the duration of 1 year).

*$P < .1$; **$P < .05$; ***$P < .01$; ****$P < .001$.

[†]Note that models 1 and 2 are based on the data obtained in part 2 and models 3 and 4 are based on the data obtained in part 2 *and* 3 of the questionnaire. In models 1 and 2, we identified speeders as respondents who completed the 2 tasks in less than 60 seconds (based on timed test of completing the tasks by 3 independent researchers) in the main analysis and in less than 25 seconds (based on the distribution of completion times; z-score ≤ -1.64) in the sensitivity analysis.

[‡]Note that the Severity and Age coefficients cannot be directly compared between the models. In models 1 and 3, these coefficients represent main effects, and in models 2 and 4, these coefficients represent conditional effects.

[§]Severity is operationalized in terms of disease-related QOL loss and measured in points from 100 on the VAS.

[‖]Age at onset of the disease (in years).

[¶]Disease outcome 1 year after falling ill.

at a decreasing marginal rate (models 3 to 5: β 1.43 to 1.46 for 20 points and β 2.82 to 2.90 for 50 points) and—compared to gains in patients who fully recovered—gains in patients who died 1 year after falling ill were, ceteris paribus, associated with a lower WTP (models 3 to 6: β -0.74 to -0.80). These results support the hypotheses that respondents' WTP is higher for more severely ill patients and larger-sized QOL gains. However, they only partially support the hypothesis that respondents' WTP is higher for younger patients because this was dependent on patients' level of disease severity in some scenarios.

The results presented in Table 5 provide further insight into the interactions between patients' disease severity, the size of QOL gains, and disease outcome and age of patients (see Fig. 1), and into the effect of respondent characteristics on WTP. The results indicate that respondents' WTP for QOL gains of different sizes in patients with different disease outcomes was dependent on patients' age. Compared to patients aged 10, respondents' WTP was lower for patients aged 70 when the QOL gain was 10 points (model 6: β -0.90) and higher for patients aged 20 when the QOL gain was 20 points (model 6: β 1.35). Compared to patients aged

Table 5. Regression results (continued).[†]

DV: WTP	Model 5[‡]		Model 6[‡]		Model 7[‡]	
	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
Severity level[§] (10 = reference)	-	-	-	-	-	-
30	0.38 (0.50)	–0.59, 1.35	0.79*** (0.28)	0.24, 1.33	0.80*** (0.28)	0.26, 1.35
50	0.71 (0.48)	–0.23, 1.64	1.19**** (0.28)	0.65, 1.74	1.20**** (0.28)	0.66, 1.75
70	0.42 (0.49)	–0.55, 1.38	1.23**** (0.26)	0.72, 1.75	1.22**** (0.26)	0.71, 1.74
90	0.08 (0.59)	–1.08, 1.24	1.34**** (0.30)	0.74, 1.93	1.35**** (0.30)	0.76, 1.95
Age[‖] (10 = reference)	-	-	-	-	-	-
20	–1.50*** (0.54)	–2.55, –0.45	–0.35 (0.24)	–0.82, 0.11	–0.17 (0.26)	–0.68, 0.33
40	–1.45*** (0.54)	–2.51, –0.38	–0.20 (0.23)	–0.66, 0.26	–0.22 (0.23)	–0.68, 0.23
70	–0.72 (0.55)	–1.80, 0.36	–0.90**** (0.24)	–1.38, –0.42	–0.99**** (0.23)	–1.44, –0.54
QOL gain (10 points = reference)	-	-	-	-	-	-
20 points	1.46**** (0.33)	0.82, 2.10	1.25*** (0.47)	0.32, 2.18	1.49**** (0.33)	0.85, 2.12
50 points	2.90**** (0.42)	2.07, 3.73	3.32**** (0.64)	2.07, 4.57	2.90**** (0.42)	2.08, 3.71
Death (Full recovery = reference)[¶]	–0.74** (0.35)	–1.44, –0.05	–0.76** (0.36)	–1.46, –0.07	–0.51 (0.51)	–1.50, 0.49
Severity × Age (Severity 10; Age 10 = reference)	-	-				
30 × 20	0.91 (0.69)	–0.43, 2.26				
30 × 40	1.35* (0.77)	–0.15, 2.85				
30 × 70	–0.67 (0.77)	–2.17, 0.83				
50 × 20	1.45** (0.69)	0.09, 2.81				
50 × 40	1.28* (0.66)	–0.02, 2.57				
50 × 70	–0.83 (0.67)	–2.13, 0.48				
70 × 20	1.65** (0.70)	0.28, 3.01				
70 × 40	1.58** (0.70)	0.20, 2.96				
70 × 70	–0.01 (0.72)	–1.42, 1.40				
90 × 20	2.39*** (0.87)	0.68, 4.09				
90 × 40	2.19*** (0.80)	0.63, 3.76				
90 × 70	0.54 (0.81)	–1.05, 2.12				
QOL gain × Age (QOL gain 10 points; Age 10 = reference)			-	-		
20 points × 20			1.35** (0.64)	0.11, 2.60		
20 points × 40			0.48 (0.57)	–0.62, 1.59		
20 points × 70			–0.88 (0.60)	–2.06, 0.29		
50 points × 20			–0.23 (0.93)	–2.06, 1.59		
50 points × 40			–0.45 (0.77)	–1.96, 1.06		
50 points × 70			–0.94 (0.74)	–2.39, 0.50		
Death × Age (Full recovery; Age 10 = reference)					-	-
Death × 20					–0.18 (0.72)	–1.60, 1.23
Death × 40					0.20 (0.65)	–1.09, 1.48
Death × 70					–1.06 (0.65)	–2.33, 0.20
Age (of respondents)	–0.35**** (0.08)	–0.51, –0.19	–0.34**** (0.08)	–0.50, –0.07	–0.34**** (0.08)	–0.50, –0.18
Age² (of respondents)	0.00**** (0.00)	0.00, 0.00	0.00**** (0.00)	0.00, 0.00	0.00**** (0.00)	0.00, 0.00
Male (Female = reference)	1.64**** (0.40)	0.85, 2.43	1.64**** (0.40)	0.86, 2.43	1.64**** (0.40)	0.86, 2.43
Children (No = reference)	1.02** (0.44)	0.16, 1.88	1.00** (0.44)	0.15, 1.86	0.99** (0.44)	0.13, 1.85
Education level[#] (Low = reference)			-	-	-	
Medium	–0.17 (0.66)	–1.46, 1.11	–0.18 (0.66)	–1.47, 1.10	–0.19 (0.66)	–1.47, 1.10
High	–0.68 (0.77)	–2.19, 0.82	–0.69 (0.77)	–2.20, 0.81	–0.69 (0.77)	–2.20, 0.82
Household income[††] (Thousands)	1.45**** (0.28)	0.90, 2.00	1.45**** (0.28)	0.90, 2.00	1.44**** (0.28)	0.89, 1.99

Table 5. Continued

DV: WTP	Model 5[‡]		Model 6[‡]		Model 7[‡]	
	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
QOL (0-100 VAS)	0.02** (0.01)	0.00, 0.04	0.02** (0.01)	0.00, 0.04	0.02** (0.01)	0.00, 0.04
CV task	−0.85**** (0.13)	−1.10, −0.60	−0.87**** (0.13)	−1.13, −0.62	−0.86**** (0.13)	−1.12, −0.61
Constant	11.89**** (2.25)	7.47, 16.30	11.27**** (2.21)	6.94, 15.61	11.26**** (2.21)	6.92, 15.60
R^2 overall	0.07		0.07		0.07	
n (observations; groups)[‡‡]	3250; 1235		3250; 1235		3250; 1235	

CV indicates contingent valuation; DV, dependent variable; NS, not stated; QOL, health-related quality of life; VAS, visual analogue scale (ranging from 0 "dead" to 100 "full health"); WTP, willingness to pay (in € per month for the duration of 1 year).

*$P < .10$; **$P < .05$; ***$P < .01$; ****$P < .001$.

[†]Note that models 5−7 are based on the data obtained in part 2 *and* 3 of the questionnaire.

[‡]Note that the Severity, Age, QOL gain, and Death coefficients cannot always be directly compared between the models. The coefficients of scenario characteristics for which no interaction effect is estimated represent main effects, and the coefficients of scenario characteristics for which an interaction effect is estimated represent conditional effects.

[§]Severity is operationalized in terms of disease-related QOL loss and measured in points from 100 on the VAS.

[ǁ]Age at onset of the disease (in years).

[¶]Disease outcome 1 year after falling ill.

[#]Low = lower vocational and primary school, Medium = middle vocational and secondary school, High = higher vocational and academic education.

[†]Household income is adjusted for household size and calculated as household income / (household size)-0.5.

[‡‡]Note that that the number of observations and groups is lower than in Table 4 as respondents who did not state their household income are excluded from Table 5.

10, respondents' WTP was lower for patients aged 70 when the patients fully recovered (model 7: β -0.99).

With regard to respondent characteristics, a higher age was associated with a lower WTP (models 4 to 6: β -0.34 to -0.35). Being male (models 5 to 7: β 1.64), having children (models 5 to 7: β 0.99 to 1.02), having a higher (adjusted) household income (models 5 to 7: β 1.44 to 1.45), and a higher QOL (models 5 to 7: β 0.02) were also associated with a higher WTP.

The results confirm the theoretical validity of the elicited WTP. The sensitivity analyses indicated that respondents' stated WTP in the practice task had a marginal effect on the stated WTP in the subsequent tasks (models 1 to 7: β 0.01) and that our results were robust.
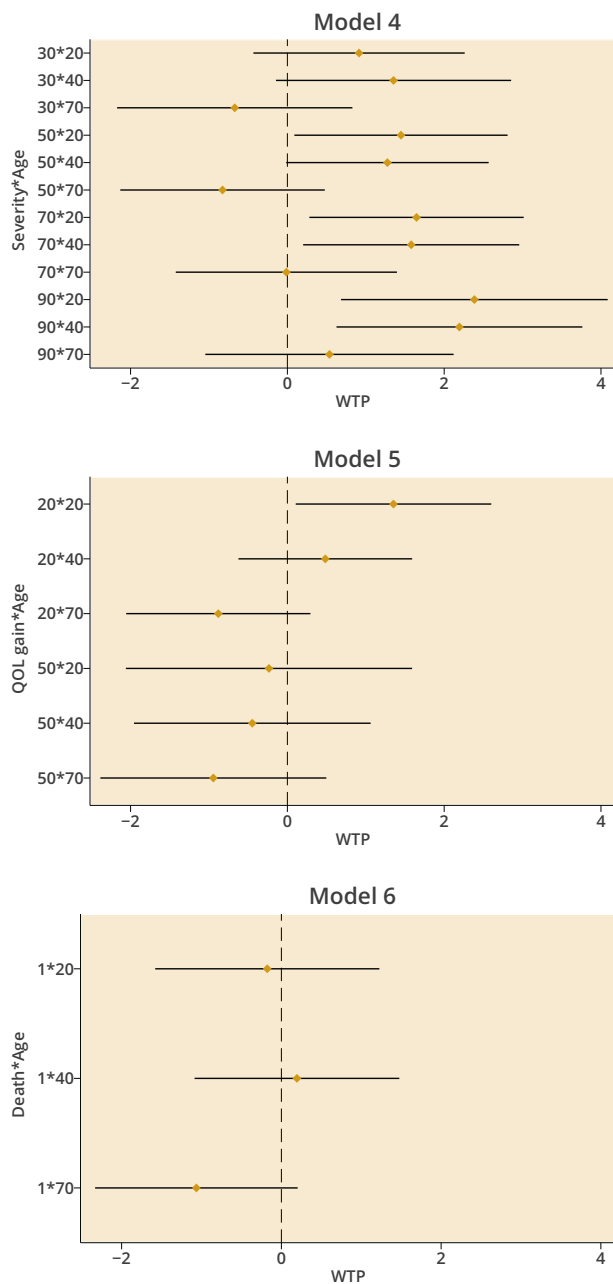
Discussion

Our aim was to examine the WTP for QOL gains in relation to the disease severity and age of patients in a representative sample of the general public in The Netherlands. Furthermore, our aim was to examine whether the WTP was different between QOL gains in patients who fully recovered and patients who died 1 year after falling ill, and whether the WTP was sensitive to scale. Our main findings are that the WTP is generally higher for QOL gains in patients with a higher level of disease severity and younger age, and for larger-sized gains, but is lower for gains in patients who die 1 year after falling ill. However, the relations were nonlinear and context dependent. For example, the WTP was higher for QOL gains in patients aged 10 than for gains in patients aged 20 and 40 when patients had a severity level of 10, but the WTP was higher for patients aged 20 and 40 from severity levels 50 and 70 onward. The WTP for QOL gains in patients aged 70 was consistently lower than for gains in younger patients. This may be explained by the fact that these patients already had their "fair share of life" at onset of the disease or that less than "full health" is more accepted at an older age.[44,55−57]

We would like to make four remarks in relation to our findings. First, we applied a SIP perspective for eliciting respondents' WTP; therefore, our findings can be driven by self-regarding as well as other-regarding preferences of respondents. Although we investigated the potential influence of observable self-regarding preferences (eg, associated with having children and respondents' proximity to the age of patients), we acknowledge that unobservable self-regarding preferences (eg, associated with the probability of respondents' own need for treatment) may have impacted our results. Second, our findings need to be considered in relation to the applied design. The WTP for QOL gains may differ when elicited on full QOL and age scales, in combination with LE gains, or from a social perspective that excludes respondents from the hypothetical patient group.[48] Third, we observed considerable preference heterogeneity, which is consistent with the findings of other studies that examined societal preferences in this context.[1,10,12,47] Accounting for (some of) this heterogeneity in resource-allocation decisions may be possible and worth pursuing, especially when aiming to align the outcomes of such decisions with societal preferences. However, our results and those of other related studies indicate that societal preferences are complex, and, consequently, there will likely always be groups in society who do (not) agree with decisions made (based on average values). Finally, the (differences in) stated WTPs could be considered modest. However, they need to be considered in relation to the respondent instruction that the increase in monthly basic health-insurance premium would apply to all adult inhabitants of The Netherlands for the duration of 1 year. Hence, on an aggregated level the (differences in) WTP per QALY can be considered substantial. The treatment generates 1000 QALYs (ie, 10 000 patients × 0.1 QALY), and hence, on average, WTP is ~1.3 million euros (calculated as 8 euros × 12 months × 13.7 million premium payers/1000 QALYs). Although this value is much higher than the monetary thresholds currently applied in The Netherlands (see Text Box 1) and likely influenced by the scenario characteristics (eg, the number of patients and certainty of QOL gains), it should be noted that such high values are not uncommon in preference-elicitation studies[58] and, considering the high ICERs of some health technologies that are currently reimbursed in The Netherlands,[59,60] also not in decision-making practice.

Our findings are consistent with those of other studies that find societal support for attaching a higher weight to health gains in more severely ill and younger patients[11,39,61−63] and to larger-sized health gains.[47] and with those of other studies that find no

Figure 1. Graphical presentation of interaction terms (mean additional effect; 95% confidence interval) presented in Table 5.

However, because these societal preferences are usually not elicited in monetary terms, we are limited in our ability to directly compare our results to those of others. However, we can compare our results to those of Bobinac et al[5] as they applied the CV approach from a social perspective in scenarios similar to ours. Both our studies found a higher WTP for health gains in younger patients and larger-sized gains. However, in contrast to our findings, Bobinac et al[5] found a lower WTP for health gains in patients with a higher level of disease severity and a higher WTP for gains in terminally ill patients. This may be explained by the different way in which they operationalized disease severity and the health gain in the specific scenarios, ie, in terms of proportional shortfall and the prevention of immediate death.[5]

The main strengths of this study lie in the use of a realistic payment vehicle, pilot-tested payment scale, and 2-step CV procedure. Although we could have applied other methods (eg, a discrete choice experiment) to elicit respondents' WTP, the CV method enabled us to approach respondents' common decision context and examine their explicit WTP (instead of, for example, deriving their WTP from the trade-off between scenario characteristics). Other strengths lie in the randomization of scenarios, exclusion of speeders, restriction of the disease duration to 1 year, standardization of patients' risk of falling ill and dying within a certain timeframe (ie, 100%, implying no uncertainty), the size of the patient group and QOL gains, and the costs of treatment, as this reduced the possible influence of an order effect, satisficing behavior,[71] cognitive biases associated with risk assessment[72] and of other considerations (eg, related to health maximization and the budget impact of reimbursing the new treatment) on our results. We appreciate that the latter strength comes with the limitation that our results cannot be generalized to other scenarios, for example, in which the number of patients is uncertain, patients are at risk of falling ill or of dying within a particular timeframe (ie, introducing uncertainty), or in which patients' lives are (also) prolonged. Another strength that comes with a limitation is the exclusion of protest zero valuations. Although including these valuations would have confounded the estimated WTP, it should be noted that the classification into protest and true zero valuations is not always straightforward and inevitably has some impact on results.[73] Some other limitations need to be discussed as well. A first limitation concerns the possible influence of payment-scale characteristics on the WTP. We facilitated a more exact mapping of respondents' WTP on the payment scale by applying a scale with a reasonable range and uneven intervals between the value points.[74,75] However, we cannot rule out the possibility that the scale influenced respondents' WTP, particularly in the case of unstable or not (yet) well-formed preferences.[74] We accounted for this by controlling for a time-effect and discussing the results in relative rather than absolute terms. A second limitation concerns the hypothetical context in which we elicited respondents' WTP.[76] Although the outlined context of a collectively funded healthcare system is realistic, we cannot rule out the possibility that their hypothetical nature increased the risk of an upward "hypothetical" bias, in which case the stated WTP could be an overestimation of respondents' true WTP.[77] However, we also cannot rule out the possibility that their realistic nature increased the risk of a downward "strategic" bias, in which case free-rider behavior of respondents may have offered a counterbalance.[78] A third limitation concerns the inclusion of QOL gains that (in some scenarios) fully restored patients' QOL to 100 points as this means we cannot distinguish between the effect of the size of QOL gains from the effect of patients' health being fully restored on respondents' WTP.[79] A final limitation concerns the low R^2 values of the regression models. We would like to note that our aim was not to predict WTP, and hence to explain as much data

support for attaching a higher weight to gains in terminally ill patients.[64–67] The latter is consistent with studies that find that the public may attach a lower weight to health gains in patients with an undesirable "end point after treatment."[58–60,68] Although there may be a moral case for attaching a higher weight to health gains in terminally ill patients,[69] our findings—like those of previous studies[64–67]—suggest that empirical support for applying a higher weight to these gains may be limited. This is recognized by the National Institute for Health and Care Excellence (NICE) in England, who recently proposed to replace their end-of-life criterion (see Text Box 1) by considerations that relate more broadly to the disease severity of patients in order to better align their decision-making framework with societal preferences.[69,70]

variance as possible. Rather, our aim was to assess whether WTP was influenced by scenario characteristics (ie, associated with the disease severity and age of patients, the size of QOL gains, and outcome of the disease) and the models successfully aided in meeting that aim. Further research is warranted to obtain insight into other factors that may influence WTP for QOL gains. This fell outside the scope of the current study.

Our results are consistent with those of other studies suggesting that equity weights based on end-of-life considerations may not be consistent with societal preferences,[64–66,80] at least not if the weight is attached to QOL gains in terminally ill patients as in the current study. Furthermore, our results are consistent with those of other studies suggesting that weights based on disease severity are consistent with societal preferences. However, our results suggest that the weights may decrease marginally with increasing disease severity, have a fairly narrow range across severity levels (possibly narrower than the threshold range of €20 000 to €80 000 currently applied in The Netherlands), and are dependent on patients' age. Further research is necessary to examine the robustness of these results in relation to the prevalence of a disease and the related budget impact of a new technology.

Because there is much variation between the results and designs of studies that examine the strength of societal preferences,[68] there is still considerable uncertainty about the "exact" weight. For example, a recent study estimated equity weights based on patients' disease severity in the range of 2.5 to 2.8 by using the person-trade-off approach.[11] Given the very limited evidence on the WTP for health gains in representative samples the of general public, further research is necessary to inform decisions about the appropriate size and range of equity weights and, in relation to this, the range and shape of the monetary thresholds against which the ICERs of new health technologies are evaluated. This may, for example, concern research into the most appropriate design for eliciting the WTP for health gains from a social or SIP perspective.

Conclusions

Our results indicate that the WTP is higher for QOL gains in more severely ill and younger patients and for larger-sized QOL gains. It is lower for QOL gains in patients who die. However, the relations are nonlinear and context dependent. These results suggest that—if the aim is to align resource-allocation decisions in healthcare with societal preferences—currently applied equity weights do not suffice.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2021.01.012.

Article and Author Information

Author Affiliations: Erasmus Centre for Health Economics Rotterdam (EsCHER), Erasmus University Rotterdam, The Netherlands (Reckers-Droog, van Exel); Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Rotterdam, The Netherlands (Reckers-Droog, van Exel, Brouwer); Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands (Reckers-Droog, van Exel, Brouwer).

Correspondence: Vivian Reckers-Droog, MSc. Erasmus School of Health Policy & Management, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands. Email: reckers@eshpm.eur.nl

REFERENCES

1. van de Wetering EJ, Stolk EA, van Exel NJA, Brouwer WBF. Balancing equity and efficiency in the Dutch basic benefits package using the principle of proportional shortfall. *Eur J Heal Econ.* 2013;14(1):107–115.
2. Weinstein MC, Torrance G, McGuire A. QALYs: the basics. *Value Health.* 2009;12:S5–S9.
3. Drummond M, Schulpher M, Claxton K, Stoddart G, Torrance G. *Methods for the Economic Evaluation of Health Care Programmes.* 4th ed. Oxford: Oxford University Press; 2015.
4. Claxton K, Paulden M, Gravelle H, Brouwer WBF, Culyer AJ. Discounting and decision making in the economic evaluation of health-care technologies. *Health Econ.* 2011;20(1):2–15.
5. Bobinac A, van de Wetering L, Van Exel NJA, Brouwer WBF. Equity-dependent social willingness to pay for a QALY. *Valuing and Refining Outcome Measures for Economic Evaluations in Health Care [dissertation].* Rotterdam: Erasmus University Rotterdam; 2016.
6. van Baal P, Perry-Duxbury M, Bakx P, Versteegh MM, van Doorslaer E, Brouwer WBF. A cost-effectiveness threshold based on the marginal returns of cardiovascular hospital spending. *Health Econ.* 2019;28(1):87–100.
7. Weinstein MC. A QALY is a QALY is a QALY—or is it? *J Health Econ.* 1988;7(3):289–290.
8. Stolk EA, Van Donselaar G, Brouwer WBF, Busschbach JJV. Reconciliation of economic concerns and health policy: illustration of an equity adjustment procedure using proportional shortfall. *Pharmacoeconomics.* 2004;22(17):1097–1107.
9. Bleichrodt H, Diecidue E, Quiggin J. Equity weights in the allocation of health care: The rank-dependent QALY model. *J Health Econ.* 2004;23(1):157–171.
10. Bobinac A, van Exel NJA, Rutten FFH, Brouwer WBF. Inquiry into the relationship between equity weights and the value of the QALY. *Value Health.* 2012;15(8):1119–1126.
11. Reckers-Droog VT, van Exel NJA, Brouwer WBF. Equity weights for priority setting in healthcare: severity, age, or both? *Value Health.* 2019;22(12):1441–1449.
12. Round J, Paulden M. Incorporating equity in economic evaluations: a multi-attribute equity state approach. *Eur J Heal Econ.* 2018;19(4):489–498.
13. Shah KK. Severity of illness and priority setting in healthcare: a review of the literature. *Health Policy.* 2009;93(2–3):77–84.
14. Schwappach DL. Resource allocation, social values and the QALY: a review of the debate and empirical evidence. *Heal Expect.* 2002;5(3):210–222.
15. Donaldson C, Baker R, Mason H, et al. The social value of a QALY: raising the bar or barring the raise? *BMC Health Serv Res.* 2011;11.
16. Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. *Health Econ.* 2005;14(2):197–208.
17. Charlton V, Rid A. Innovation as a value in healthcare priority-setting: the UK experience. *Soc Justice Res.* 2019;32(2):208–238.
18. Olsen JA, Richardson J, Dolan P, Menzel P. The moral relevance of personal characteristics in setting health care priorities. *Soc Sci Med.* 2003;57(7):1163–1172.

19. van Exel NJA, Baker R, Mason H, Donaldson C, Brouwer WBF. Public views on principles for health care priority setting: findings of a European cross-country study using Q methodology. *Soc Sci Med*. 2015;126:128–137.

20. Norheim OF, Baltussen R, Johri M, et al. Guidance on priority setting in health care (GPS-Health): the inclusion of equity criteria not captured by cost-effectiveness analysis. *Cost Eff Resour Alloc*. 2014;12(1):1–8.

21. Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Econ*. 1999;8(1):25–39.

22. Accountability Daniels. for reasonableness: establishing a fair process for priority setting is easier than agreeing on principles. *Acad Manag Rev*. 2000;31(2):386–408.

23. Maynard A. European health policy challenges. *Health Econ*. 2005;14(Suppl 1):255–263.

24. Daniels N, Sabin JE. Accountability for reasonableness: an update. *BMJ*. 2008;337.

25. Chalkidou K, Glassman A, Marten R, et al. Priority-setting for achieving universal health coverage. *Bull World Health Organ*. 2016;94(6):462–467.

26. Broqvist M, Sandman L, Garpenby P, Krevers B. The meaning of severity: do citizens' views correspond to a severity framework based on ethical principles for priority setting? *Health Policy*. 2018;122(6):630–637.

27. Le Polain M, Franken M, Koopmanschap M, Cleemput I. Drug reimbursement systems: International comparison and policy recommendations (KCE Report C147). *Health Services Research*. 2010.

28. Zorginstituut Nederland. *Ziektelast in de praktijk*. Diemen; 2018.

29. Zorginstituut Nederland. *Cost-effectiveness in practice*. Diemen; 2015.

30. National Institute for Health and Care Excellence. *Guide to the methods of technology appraisal*. London. 2013.

31. Norwegian Ministry of Health and Care Services. *Principles for priority setting in health care*. Oslo. 2017.

32. Magnussen J, Aaserud M, Granaas T, Al E. På ramme alvor—Alvorlighet og prioritering. https://www.regjeringen.no/contentassets/d5da48ca5d1a4b128 c72fc5daa3b4fd8/paa_ramme_alvor.pdf.

33. Magnussen J. Severity of illness and priority setting in Norway: summary of a report from a working group. https://www.regjeringen.no/contentassets/d5 da48ca5d1a4b128c72fc5daa3b4fd8/summary_the_magnussen_report_on_ severity.pdf. Accessed January 13, 2020.

34. Ottersen T, Førde R, Kakad M, et al. A new proposal for priority setting in Norway: open and fair. *Health Policy*. 2016;120(3):246–251.

35. Statens Legemiddelverk. Beregning av absolutt prognosetap i hurtig metodevurderinger av legemidler – bruke beregningsprinsippet for behandling eller forebygging? 2019. https://legemiddelverket.no/Documents/Offentlig finansiering og pris/Dokumentasjon til metodevurdering/APT- behandling eller forebygging_310119.pdf.

36. Zorginstituut Nederland. *Uitvoeringstoets lage-ziektelastbenadering*. Diemen; 2012.

37. Reckers-Droog VT, van Exel NJA, Brouwer WBF. Looking back and moving forward: on the application of proportional shortfall in healthcare priority setting in The Netherlands. *Health Policy*. 2018;122(6):621–629.

38. Drummond M. Twenty years of using economic evaluations for drug reimbursement decisions: what has been achieved? *J Health Polit Policy Law*. 2013;38(6):1081–1102.

39. Dolan P, Tsuchiya A. Health priorities and public preferences: The relative importance of past health experience and future health prospects. *J Health Econ*. 2005;24(4):703–714.

40. Stolk EA, Pickee SJ, Ament AHJA, Busschbach JJV. Equity in health care prioritisation: an empirical inquiry into social value. *Health Policy*. 2005;74(3):343–355.

41. Olsen JA. Priority preferences: "end of life" does not matter, but total life does. *Value Health*. 2013;16(6):1063–1066.

42. Richardson J, Iezzi A, Maxwell A, Chen G. Does the use of the proportional shortfall help align the prioritisation of health services with public preferences? *Eur J Heal Econ*. 2018;19(6):797–806.

43. Cookson R. Can the nice "end-of-life premium" be given a coherent ethical justification? *J Health Polit Policy Law*. 2013;38(6):1129–1148.

44. Williams A. Intergenerational equity: an exploration of the "fair innings" argument. *Health Econ*. 1997;6(2):117–132.

45. Collins M, Latimer N. NICE's end of life decision making scheme: impact on population health. *BMJ*. 2013;346(7905):1–5.

46. Cancer Research United Kingdom. Cancer mortality by age. October 10, 2019. https://www.cancerresearchuk.org/health-professional/cancer-statistics/mo rtality/age#heading-Zero. Accessed April 1, 2020.

47. Bobinac A, van Exel NJA, Rutten FFH, Brouwer WBF. Get more, pay more? An elaborate test of construct validity of willingness to pay per QALY estimates obtained through contingent valuation. *J Health Econ*. 2012;31(1):158–168.

48. Dolan P, Olsen JA, Menzel P, Richardson J. An inquiry into the different perspectives that can be used when eliciting preferences in health. *Health Econ*. 2003;12(7):545–551.

49. Bobinac A, Van Exel NJA, Rutten FFH, Brouwer WBF. Valuing qaly gains by applying a societal perspective. *Health Econ*. 2013;22(10):1272–1281.

50. Bobinac A, van Exel NJA, Rutten FFH, Brouwer WBF. Willingness to pay for a quality-adjusted life-year: the individual perspective. *Value Health*. 2010;13(8):1046–1055.

51. Ministry of Health W and S. Wat kost het om een zorgverzekering af te sluiten?. https://www.rijksoverheid.nl/onderwerpen/zorgverzekering/vraag-en-antwoord/premie-zorgverzekering. Accessed January 13, 2020.

52. Johnson DS, Smeeding TM, Torrey BB. Economic inequality through the prisms of income and consumption. *Mon Labor Rev*. 2005;128(4):11–24.

53. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science*. 1981;211(4481):453–458.

54. Abasolo I, Tsuchiya A. Exploring social welfare functions and violation of monotonicity: an example from inequalities in health. *J Health Econ*. 2004;23(2):313–329.

55. Wouters S, van Exel NJA, Rohde KIM, Vromen JJ, Brouwer WBF. Acceptable health and priority weighting: discussing a reference-level approach using sufficientarian reasoning. *Soc Sci Med*. 2017;181:158–167.

56. Wouters S, van Exel NJA, Rohde KIM, Brouwer WBF. Are all health gains equally important? An exploration of acceptable health as a reference point in health care priority setting. *Health Qual Life Outcomes*. 2015;13(1):1–10.

57. Zrubka Z, Hermann Z, Gulácsi L, Brodszky V, Rencz F, Péntek M. Determinants of the acceptability of health problems in different ages: exploring a new application of the EQ VAS. *Eur J Heal Econ*. 2019;20(0123456789):31–41.

58. Baker R, Bateman I, Donaldson C, Jones-Lee M, Lancsar E, Loomes G, et al. Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the social value of a QALY project. *Health Technol Assess*. 2010;14(27).

59. Rombach S, Hollak C, Linthorst G, Dijkgraaf M. Enzyme replacement therapy for Fabry disease. *Enzym Technol Pluripotent Play Discov Ther Agent*. 2013;321–334.

60. Kanters TA, van der Ploeg AT, Kruijshaar ME, et al. Cost-effectiveness of enzyme replacement therapy with alglucosidase alfa in adult patients with Pompe disease. *Orphanet J Rare Dis*. 2017;12(1):1–8.

61. Richardson J, Iezzi A, Maxwell A. How important is severity for the evaluation of health services: new evidence using the relative social willingness to pay instrument. *Eur J Heal Econ*. 2017;18(6):671–683.

62. Richardson J, McKie J, Iezzi A, Maxwell A. Age weights for health services derived from the relative social willingness-to-pay instrument. *Med Decis Mak*. 2017;37(3):239–251.

63. Shiroiwa T, Igarashi A, Fukuda T, Ikeda S. WTP for a QALY and health states: more money for severer health states? *Cost Eff Resour Alloc*. 2013;11(1):1–7.

64. McHugh N, Pinto-Prades JL, Baker R, Mason H, Donaldson C. Exploring the relative value of end of life QALYs: are the comparators important? *Soc Sci Med*. 2020;245:112660.

65. Shah KK, Tsuchiya A, Wailoo AJ. Valuing health at the end of life: a review of stated preference studies in the social sciences literature. *Soc Sci Med*. 2018;204:39–50.

66. Gyrd-Hansen D. Is there additional value attached to health gains at the end of life? A revisit. *Health Econ*. 2018;27(1):e71–e75.

67. Hansen LD, Kjær T. Disentangling public preferences for health gains at end-of-life: further evidence of no support of an end-of-life premium. *Soc Sci Med*. 2019;236:112375.

68. Nord E, Johansen R. Concerns for severity in priority setting in health care: a review of trade-off data in preference studies and implications for societal willingness to pay for a QALY. *Health Policy*. 2014;116(2–3):281–288.

69. National Institute for Health and Care Excellence. *Modifiers: Task and finish group report*. London. 2020.

70. National Institute for Health and Care Excellence. *The NICE methods of health technology evaluation: the case for change*. London. 2020.

71. Zhang C, Conrad FG. Speeding in web surveys: the tendency to answer very fast and its association with straightlining. *Surv Res Methods*. 2014;8(2):127–135.

72. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*. 1974;185(4157):1124–1131.

73. Rankin J, Robinson A. *Accounting for protest zeros in contingent valuation: a review of literature*. HEG Work Paper; 2018, 18–01.

74. Soeteman L, van Exel J, Bobinac A. The impact of the design of payment scales on the willingness to pay for health gains. *Eur J Heal Econ*. 2017;18(6):743–760.

75. Whynes DK, Sach TH. WTP and WTA: do people think differently? *Soc Sci Med*. 2007;65(5):946–957.

76. Bobinac A. Mitigating hypothetical bias in willingness to pay studies: post-estimation uncertainty and anchoring on irrelevant information. *Eur J Heal Econ*. 2019;20(1):75–82.

77. Poder TG, He J. Willingness to pay and the sensitivity of willingness to pay for interdisciplinary musculoskeletal clinics: a contingent valuation study in Quebec, Canada. *Int J Heal Econ Manag*. 2016;16(4):337–3361.

78. Johannesson M, Jönsson B. Economic evaluation in health care: is there a role for cost-benefit analysis? *Health Policy*. 1991;17(1):1–23.

79. Shah KK, Tsuchiya A, Wailoo AJ. Valuing health at the end of life: an empirical study of public preferences. *Eur J Heal Econ*. 2014;15(4):389–399.

80. Lancsar E, Gu Y, Gyrd-Hansen D, et al. The relative value of different QALY types. *J Health Econ*. 2020;70:102303.