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

Equity Weights for Priority Setting in Healthcare: Severity, Age, or Both?

Vivian Reckers-Droog, MSc,[1,*] Job van Exel, PhD,[1,2] Werner Brouwer, PhD[1,2]

[1]Erasmus University Rotterdam, Erasmus School of Health Policy & Management, Rotterdam, the Netherlands; [2]Erasmus University Rotterdam, Erasmus School of Economics, Rotterdam, the Netherlands.

ABSTRACT

Background: Priority setting in healthcare can be guided by both efficiency and equity principles. The latter principle is often explicated in terms of disease severity and, for example, defined as absolute or proportional shortfall. These severity operationalizations do not explicitly consider patients' age, even though age may be inextricably related to severity and an equity-relevant characteristic.

Objective: This study examines the relative strength of societal preferences for severity and age for informing allocation decisions in healthcare.

Methods: We elicited preferences for severity and age in a representative sample of the public in the Netherlands (N = 1025) by applying choice tasks and person-trade-off tasks in a design in which severity levels and ages varied both separately and simultaneously between patient groups. We calculated person trade-off ratios and, in addition, applied ordinary least squares regression models to aid interpretation of the ratios when both severity and age varied.

Results: Respondents attached a higher weight (median of ratios: 2.46-3.50) to reimbursing treatment for relatively more severely ill and younger patients when preferences for both were elicited separately. When preferences were elicited simultaneously, respondents attached a higher weight (median of ratios: 1.98 and 2.42) to reimbursing treatment for relatively younger patients, irrespective of patients' severity levels. Ratios varied depending on severity level and age and were generally higher when the difference in severity and age was larger between groups.

Conclusions: Our results suggest that severity operationalizations and equity weights based on severity alone may not align with societal preferences. Adjusting decision-making frameworks to reflect age-related societal preferences should be considered.

Keywords: age, equity, priority setting, person trade-off, severity of illness.

VALUE HEALTH. 2019; ■(■):■–■

Introduction

The increasing demand for healthcare and the resulting pressure on scarce resources render healthcare priority setting inevitable. Allocation decisions can be guided by both efficiency and equity principles, and hence be informed by the cost-effectiveness of health technologies and equity considerations associated with, for example, patient and disease characteristics.[1-6] Although it has been advocated to explicitly and transparently incorporate these principles into the decision-making framework,[7-9] there is little agreement on how equity considerations should be defined and how the trade-off between efficiency and equity, or between equity considerations, should be operationalized.[3,5] This may partly explain why, so far, only a few countries integrated equity considerations and weights into formal decision-making frameworks.[3] Examples of countries that have done so are Norway and the Netherlands.[3,10-21] Text box 1 includes a brief overview of the equity considerations and decision-making frameworks applied in these 2 countries.

Although the decision-making frameworks applied in Norway and the Netherlands both account for societal preferences relating to the disease severity of patients, neither explicitly accounts for age-related societal preferences. This may be (partially) explained by the fact that, politically, age usually is not regarded as a relevant or even acceptable decision criterion.[22-24] It should, however, be noted that—in decision-making practice—severity is not independent from age, and not explicitly accounting for age in allocation decisions can still result in a prioritization that favors certain age groups over others.[25] For example, in Norway, younger patients are, ceteris paribus, more likely to lose a larger absolute amount of their remaining

* Address correspondence to: Vivian Reckers-Droog, MSc, Erasmus University Rotterdam, Erasmus School of Health Policy & Management, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands. Email: reckers@eshpm.eur.nl

https://doi.org/10.1016/j.jval.2019.07.012

BOX 1. Equity considerations and decision-making frameworks in Norway and the Netherlands.

In Norway, equity considerations are explicated in terms of disease severity and defined as absolute shortfall (AS). AS is operationalized as the disease-related loss of remaining quality-adjusted life-years (QALYs) without the new technology, compared to the QALY expectation in absence of the disease.[10-12] The Magnussen Committee suggested to divide AS into 6 severity classes and evaluate the incremental cost-effectiveness ratios (ICERs) of health technologies that target diseases with an AS of 0 to 0.39, 4 to 7.9, 8 to 11.9, 12 to 15.9, 16 to 19.9, and >20 QALYs against thresholds of NOK 275 000/QALY, NOK 385 000/QALY, NOK 495 000/QALY, NOK 605 000/QALY, NOK 715 000/QALY, and NOK 825 000/QALY, respectively (range approximately €27 500–€82 500).[13,14] Although the Norwegian government did not formally adopt the suggested AS classes and thresholds, they are informally used to inform allocation decisions in healthcare.[11]

In the Netherlands, equity considerations are also explicated in terms of severity, and severity is defined in terms of proportional shortfall (PS). PS is operationalized as the fraction of disease-related QALY loss without the new technology, relative to the remaining QALY expectation in absence of the disease and measured on a scale from 0 "no QALY loss" to 1 "complete loss of remaining QALYs".[15,16-18] PS is divided into 4 severity classes. Health technologies that target diseases with a PS level of <0.10 are, in principle, not reimbursed.[15,16] The ICERs of health technologies that target diseases with PS levels of 0.10 to 0.40, 0.41 to 0.70, and 0.71 to 1.00 are evaluated against reference values of €20 000/QALY, €50 000/QALY, and €80 000/QALY, respectively.[15,16,18-20] Since 2018, the Dutch National Health Care Institute (ZIN) supplements information on PS with information on patients' AS and prospective health (PH) to be transparent about the possible consequences of applying PS for different age groups.[15] PH is a severity operationalization that considers patients' expected health and death and prioritizes those with the worst prognosis without the new technology.[21] PH is calculated as the remaining QALY expectation from the onset of disease in case of no treatment.[21]

quality-adjusted life-years (QALYs), and hence may indirectly be prioritized over older patients. Moreover, for older patients it becomes increasingly difficult, if not impossible, to fall into the highest severity class. For example, 65-year-old patients may, on average, have no more than 15 QALYs left to lose. In the Netherlands, proportional shortfall (PS) was implemented with the intention to avoid ageism in allocation decisions by enabling patients of all ages to lose the same relative amount of their remaining QALYs.[15,18] However, older patients are, ceteris paribus, more likely to lose a larger fraction of their remaining QALYs, and hence may indirectly be prioritized over younger patients.[12]

Although equity weights based on absolute shortfall (AS) and PS may be inextricably related to patients' age, neither explicitly distinguishes between age groups nor aims to weight age in allocation decisions. Hence, prioritizing between age groups with AS or PS levels that are equal or fall into the same severity class is not possible within current decision-making frameworks. Empirical evidence, however, suggests that the public considers information on severity *and* age important and generally prefers prioritizing younger over older patients.[6,17,26-33,35,36] At the same time, evidence regarding the direct trade-off between preferences for severity and age and the relative weight these should receive in allocation decisions is limited.[17,26] The aim of this study is therefore to contribute to the existing literature on this topic by examining the relative strength of societal preferences for severity and age. To meet this aim, we applied the person-trade-off (PTO) approach in a design in which severity levels and ages varied both separately and simultaneously between patient groups. We compared preferences for severity and age in relation to AS, PS, and prospective health (PH), as these severity operationalizations are currently applied in Norway and the Netherlands. The results may be of interest to all countries seeking to understand equity considerations associated with severity and age and to operationalize these considerations in the context of the healthcare priority setting.

Methods

Sample and Data Collection

A questionnaire was designed and subsequently distributed online in October 2018. Respondents (N = 1025) were quota sampled to represent the public in the Netherlands in terms of age (18-70 years), sex, and education level (Table 1). Before collecting the data, we pilot tested the comprehensiveness of the questionnaire and clarity of the applied concepts and choice and PTO tasks in 2 consecutive samples (n = 120 and 1023).

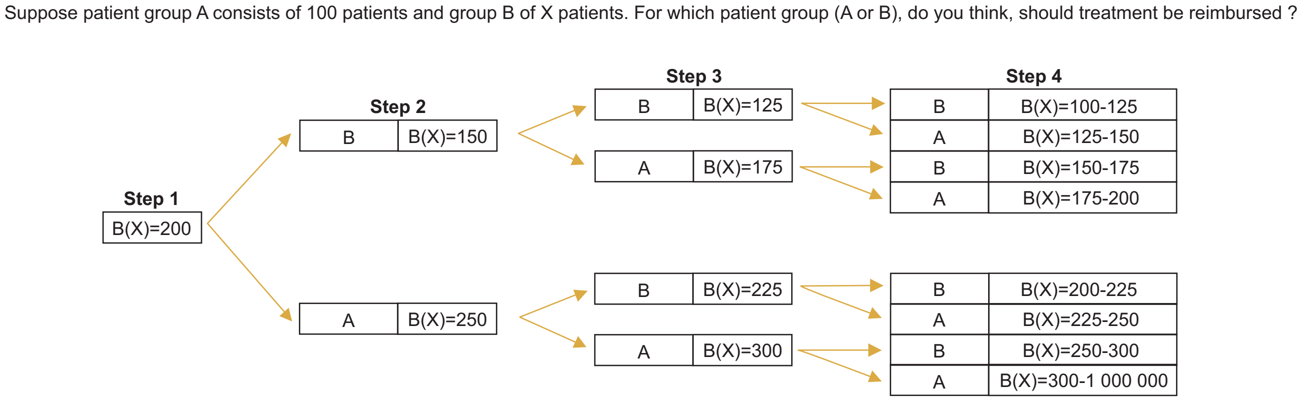
Before respondents completed the questionnaire, we explained that healthcare resources are scarce and policy makers inevitably have to choose between competing health technologies and patient groups for reimbursement. We asked respondents to advise policy makers faced with the choice of reimbursing a health technology for 1 of 2 patient groups (labeled "A" and "B") on how best to allocate the available budget. We explained that there were no differences between the patient groups other than those explicated and that the treatment type and costs were the same for both groups. We limited the duration of the disease and treatment-related health gain to 1 year to standardize the health gains from treatment in the different age groups. This avoided the influence of other considerations, including preferences for lifetime health and health maximization. Based on this temporary loss in health-related quality of life (QOL) and no loss in length of life, we calculated patients' AS, PS, and PH (not shown to respondents)

Table 1. Sample characteristics (N = 1014).*

	%	Mean (SD)
Age (y)		45.3 (14.7)
Sex (female)	51.7	
Education level[†]		
Low	8.3	
Medium	42.3	
High	49.4	
Health status (in points on VAS)		75.4 (17.7)
Completion time of PTO tasks (min)		7.7 (5.1)

PTO indicates person-trade-off; VAS, visual analogue scale (range 0-100).
*Respondents who completed the PTO tasks in less than 115 seconds (n = 11) are excluded from this table.
[†]Low = lower vocational and primary school; medium = middle vocational and secondary school; high = higher vocational and academic education.

Figure 1. Graphical representation of the iterative person trade-off process.



Note that in step 4, respondents indicated the number of patients within the indicated range. For respondents with a consistent preference for one of the patient groups, this range was restricted to a maximum of 1 000 000 patients.

and compared preferences for severity and age in relation to these severity operationalizations.

Questionnaire

The questionnaire consisted of 3 parts. In part 1, we introduced 3 central concepts in text and graphs to respondents: (1) QOL, operationalized in points on a visual analogue scale (VAS) that ranged from 0 "the worst health you can imagine" to 100 "the best health you can imagine"[34]; (2) severity, operationalized in terms of disease-related QOL loss; and (3) treatment-related QOL gain. To familiarize respondents with these concepts and the applied tasks, we asked them to rate their current health on the VAS and complete 1 example choice task and 1 example PTO task. We further asked them to assess the clarity of the concepts and example tasks on a 7-point Likert scale ranging from 1 "very unclear" to 7 "very clear".

In part 2, respondents stated their preferences for severity and age by completing a total of 6 randomly assigned choice tasks and 6 associated PTO tasks in 3 similarly structured modules. In the choice tasks, respondents stated their preference for reimbursing treatment for either patient group A or B, or their indifference between the groups, based on the contingency that both groups consisted of 100 patients. In the associated PTO tasks, we set the respondents' group of preference (or, in case of indifference, a randomly selected group) as a reference group consisting of 100 patients. We asked respondents to indicate in 4 iterative steps of how many patients (between 100 and 1 000 000) the other group should consist to be indifferent between the 2 groups. Figure 1 presents a graphical representation of the PTO process, including the iterative steps taken to elicit respondents' preferences and the intervals between the steps. A detailed description of the applied PTO approach and examples of the choice and PTO tasks are included in the Appendix in the Supplemental Materials found at https://doi.org/10.1016/j.jval.2019.07.012.

We started each module with introducing the patient groups and explaining that the groups would have lived in full health (a score of 100 on the VAS) until the age of 80 if they had not fallen ill at the age of *X* (with *X* = 10, 40, or 70). We then explained that the disease affected patients' QOL but not their life expectancy. Because of the disease, their QOL decreased from 100 to *Y* (with *Y* = 20, 50, or 80) on the VAS for the duration of 1 year, after which it restored to the initial level. If patients would receive treatment, their QOL would increase with 20 points during that year. In module 1, respondents completed 2 choice tasks and 2 associated PTO tasks based on a "small" and "large" difference in severity between the patient groups (ie, 30 points, from 100 down to 20 vs 50 or 50 vs 80, and 60 points, from 100 down to 20 vs 80, respectively), while patients' ages were kept constant. In module 2, respondents completed 2 choice tasks and 2 associated PTO tasks based on a "small" and "large" difference in age between the groups (ie, 30 years, age 10 vs 40 or 40 vs 70, and 60 years, age 10 vs 70, respectively), while patients' severity levels were kept constant. In module 3, respondents completed 2 choice tasks and 2 associated PTO tasks based on a "small" and "large" difference in both severity and age between the groups (ie, 30 points/years and 60 points/years, respectively).

Table 2 presents the applied attributes and levels, and Table 3 in the Results section includes the applied choice sets and associated AS, PS, and PH levels. We reduced the possible risk of order effects by presenting modules 1 and 2, the choice sets, and patient groups in random order to respondents.

In part 3 of the questionnaire, we asked respondents about their socio-demographic characteristics.

Statistical Analyses and Hypotheses

We examined preferences for severity and age by calculating the percentage of total with a preference for reimbursing treatment for patient group A, group B, or neither of the groups and

Table 2. Overview of attributes and levels.

Attribute	Levels
Severity* (in points on VAS)	20; 50; 80
Age (y)	10; 40; 70
QOL after treatment† (in points on VAS)	40; 70; 100
Treatment-related QOL gain (in points on VAS)	20
Life expectancy (y)	80

QOL indicates health-related quality of life; VAS, visual analogue scale (range 0-100).
*Severity is operationalized in terms of disease-related QOL loss and measured in points from 100 on the VAS.
†QOL after treatment is calculated as 100 − disease-related QOL loss + treatment-related QOL gain.

Table 3. Preferences for reimbursing treatment for patient group A, group B, or neither of the groups (in % of total; N = 1014).*

	CS	Severity[†]		Age		Absolute shortfall[‡]		Proportional shortfall[§]		Prospective health[‖]		n	Preference (in %)		
		A	B	A	B	A	B	A	B	A	B		A	B	Neither
Module 1	1	80	10	10	10	0.8[¶]	0.5	0.01[¶]	0.01	69.2[¶]	69.5	165	50.9	24.2	24.9
	2	50	20	10	10	0.5[¶]	0.2	0.01[¶]	0.00	69.5[¶]	69.8	171	58.5	19.9	21.6
	3	80	50	40	40	0.8[¶]	0.5	0.02[¶]	0.01	39.2[¶]	39.5	173	47.4	32.4	20.2
	4	50	20	40	40	0.5[¶]	0.2	0.01[¶]	0.01	39.5[¶]	39.8	170	57.6	20.6	21.8
	5	80	50	70	70	0.8[¶]	0.5	0.08[¶]	0.05	9.2[¶]	9.5	165	48.5	27.9	23.6
	6	50	20	70	70	0.5[¶]	0.2	0.05[¶]	0.02	9.5[¶]	9.8	170	58.8	21.8	19.4
	7	80	20	10	10	0.8[¶]	0.2	0.01[¶]	0.00	69.2[¶]	69.8	336	58.9	21.7	19.4
	8	80	20	40	40	0.8[¶]	0.2	0.02[¶]	0.01	39.2[¶]	39.8	338	57.4	25.7	16.9
	9	80	20	70	70	0.8[¶]	0.2	0.08[¶]	0.02	9.2[¶]	9.8	340	54.1	23.8	22.1
Module 2	1	80	80	10	40	0.8	0.8	0.01	0.02	69.2	39.2	161	52.8	14.9	32.3
	2	80	80	40	70	0.8	0.8	0.02	0.08	39.2	9.2	171	60.8	8.2	31.0
	3	50	50	10	40	0.5	0.5	0.01	0.01	69.5	39.5	170	52.4	13.5	34.1
	4	50	50	40	70	0.5	0.5	0.01	0.05	39.5	9.5	169	55.6	13.6	30.8
	5	20	20	10	40	0.2	0.2	0.00	0.01	69.8	39.8	171	44.4	20.5	35.1
	6	20	20	40	70	0.2	0.2	0.01	0.02	39.8	9.8	172	55.2	20.4	24.4
	7	80	80	10	70	0.8	0.8	0.01	0.08	69.2	9.2	338	61.8	13.9	24.3
	8	50	50	10	70	0.5	0.5	0.01	0.05	69.5	9.5	341	55.7	16.1	28.2
	9	20	20	10	70	0.2	0.2	0.00	0.02	69.8	9.8	335	63.3	11.3	25.4
Module 3	1	80	50	10	40	0.8[¶]	0.5	0.01	0.01	69.2	39.5	129	57.4	19.4	23.2
	2	80	50	40	10	0.8	0.5	0.02	0.01	39.2	69.5	130	30.8	38.4	30.8
	3	80	50	40	70	0.8[¶]	0.5	0.02	0.05	39.2	9.5	129	60.5	20.1	19.4
	4	80	50	70	40	0.8	0.5	0.08	0.01	9.2	39.5	129	25.6	48.1	26.3
	5	50	20	10	40	0.5[¶]	0.2	0.01[¶]	0.01	69.5	39.8	127	74.0	12.6	13.4
	6	50	20	40	10	0.5[¶]	0.2	0.01[¶]	0.00	39.5[¶]	69.8	123	45.5	30.9	23.6
	7	50	20	40	70	0.5[¶]	0.2	0.01	0.02	39.5	9.8	125	68.8	11.2	20.0
	8	50	20	70	40	0.5	0.2	0.05	0.01	9.5	39.8	122	38.5	41.8	19.7
	9	80	20	10	70	0.8[¶]	0.2	0.01	0.02	69.2	9.8	514	70.4	13.0	16.6
	10	80	20	70	10	0.8	0.2	0.08	0.00	9.2	69.8	500	36.0	44.6	19.4

A indicates patient group A; B, patient group B; CS, choice set; QOL, health-related quality of life.
*Respondents who completed the person-trade-off tasks in less than 115 seconds (n = 11) are excluded from this table.
[†]Severity is operationalized in terms of disease-related QOL loss and measured in points from 100 on the 0-100 visual analogue scale.
[‡]Absolute shortfall is calculated as ([disease-related QOL loss/100] × 1 year).
[§]Proportional shortfall is calculated as ([disease-related QOL loss/100] × 1 year)/(80 - age at onset of disease).
[‖]Prospective health is calculated as 80 - age at onset of disease-absolute shortfall.
[¶]Choice set in which preferences are aligned with absolute shortfall, proportional shortfall, or prospective health.

the mean- and median-based PTO ratios. A PTO ratio of 1 indicates that respondents, on average, attach an equal weight to reimbursing treatment for groups A and B, and a ratio of >1 (<1) that they attach a higher (lower) weight to reimbursing treatment for group A. We present the unpooled (ie, per choice set) and pooled (ie, per "small" and "large" difference) mean- and median-based PTO ratios, including our calculation methods, in Table 4. We focus the discussion of the results on the ratio of medians (ROMs) and median of ratios (MORs) to account for outliers.

To aid interpretation of the PTO ratios in module 3, we applied ordinary least squares (OLS) regression models in which we assumed utility equivalence between reimbursing treatment for the patient groups. We regressed the difference in patient numbers (multiplied by patients' treatment-related QOL gain) on the differences in severity level (ΔSeverity) and age (ΔAge) between the groups and calculated the marginal rate of substitution (MRS; 95% confidence interval [CI]) between ΔSeverity and ΔAge by bootstrapping (5000 repetitions) the ratio of the coefficients. We present the OLS results both including and excluding respondents with outlying preferences ($-1.96 >$ z-score > 1.96) in Table 5. We focus the discussion of the results on the latter to account for outlying preferences. Note that, for calculating the preferences (in percentages) and PTO ratios, we restructured the patient-group order to A vs B for all respondents. For the regression models, we used the order in which the groups were presented to respondents (A vs B or

Table 4. PTO results (N = 1014).*

		CS	Severity†		Age		n	Mean raw responses		Ratio of means	Median raw responses		Ratio of medians (ROMs)
			A	B	A	B		M_A	M_B	M_B / M_A	$\tilde{M}_A$	$\tilde{M}_B$	$\tilde{M}_B / \tilde{M}_A$
Module 1		1	80	50	10	10	165	217.48	9764.52	44.90	100.00	160.00	1.60
		2	50	20	10	10	171	322.12	10 190.43	31.64	100.00	251.00	2.51
		3	80	50	40	40	173	836.27	3107.00	3.72	100.00	130.00	1.30
		4	50	20	40	40	170	227.22	8965.28	39.46	100.00	275.00	2.75
		5	80	50	70	70	165	928.43	10 622.85	11.44	100.00	160.00	1.60
		6	50	20	70	70	170	1059.85	2918.43	2.75	100.00	289.50	2.90
	N						1014						
	Ratio of means							599.24	7558.40	12.61			
	Median of ratios (MORs)												2.51
		7	80	20	10	10	336	545.00	13 803.70	25.33	100.00	301.00	3.01
		8	80	20	40	40	338	276.71	2815.54	10.18	100.00	275.00	2.75
		9	80	20	70	70	340	687.43	12 814.76	18.64	100.00	275.00	2.75
	N						1014						
	Ratio of means							503.33	9809.38	19.49			
	Median of ratios (MORs)												2.75
Module 2		1	80	80	10	40	161	200.21	3580.17	17.88	100.00	224.00	2.24
		2	80	80	40	70	171	133.35	9672.78	72.54	100.00	275.00	2.75
		3	50	50	10	40	170	253.68	2883.51	11.37	100.00	210.00	2.10
		4	50	50	40	70	169	184.37	15 797.84	85.69	100.00	301.00	3.01
		5	20	20	10	40	171	224.25	11 189.68	49.90	100.00	210.00	2.10
		6	20	20	40	70	172	871.80	2146.66	2.46	100.00	267.50	2.68
	N						1014						
	Ratio of means							313.23	7567.21	24.16			
	Median of ratios (MORs)												2.46
		7	80	80	10	70	338	128.17	12 564.78	98.03	100.00	400.00	4.00
		8	50	50	10	70	341	474.91	13 564.93	28.56	100.00	350.00	3.50
		9	20	20	10	70	335	126.67	4900.60	38.69	100.00	350.00	3.50
	N						1014						
	Ratio of means							244.28	10 369.07	42.45			
	Median of ratios (MORs)												3.50
Module 3		1	80	50	10	40	129	224.66	18 543.20	82.54	100.00	275.00	2.75
		2	80	50	40	10	130	1234.77	1205.29	0.98	112.50	100.00	0.89
		3	80	50	40	70	129	140.09	18 591.40	132.71	100.00	301.00	3.01
		4	80	50	70	40	129	3519.44	1009.33	0.29	140.00	100.00	0.71
		5	50	20	10	40	127	121.50	4773.40	39.29	100.00	400.00	4.00
		6	50	20	40	10	123	2000.12	18 575.50	9.29	100.00	120.00	1.20
		7	50	20	40	70	125	920.50	7317.10	7.95	100.00	350.00	3.50
		8	50	20	70	40	120	1254.18	2552.46	2.04	112.50	100.00	0.89
	N						1014						
	Ratio of means							1174.65	9067.36	7.72			

continued on next page

Table 4. Continued

	CS	Severity†		Age		n	Mean raw responses		Ratio of means	Median raw responses		Ratio of medians (ROMs)
		A	B	A	B		M_A	M_B	M_B / M_A	$\tilde{M}_A$	$\tilde{M}_B$	$\tilde{M}_B / \tilde{M}_A$
Median of ratios (MORs)												1.98
	9	80	20	10	70	514	176.76	12976.93	73.42	100.00	400.00	4.00
	10	80	20	70	10	500	5325.84	5818.83	1.09	120.00	100.00	0.83
N						1014						
Ratio of means							2715.76	9447.29	3.48			
Median of ratios (MORs)												2.42

A indicates patient group A; B, patient group B; CS, choice set; PTO, person trade-off; QOL, health-related quality of life.
*Respondents who completed the PTO tasks in less than 115 seconds (n = 11) are excluded from this table.
†Severity is operationalized in terms of disease-related QOL loss and measured in points from 100 on the 0-100 visual analogue scale.

B vs A). Although this was necessary for conducting the different analyses, the difference in data structure and applied methods hampers a direct comparison between the PTO ratios and the ratios obtained from the OLS regressions.

We determined a minimum completion time of 115 seconds for the PTO tasks based on the distribution of the time respondents needed for completing the tasks in the pilot study and a timed test of the tasks by 3 independent researchers. We excluded respondents who completed the tasks more quickly from the analyses as we expected these "speeders" to have completed the tasks too quickly to properly read, understand, and complete them. We performed sensitivity analyses to test the robustness of our results by alternately repeating the analyses excluding respondents who reported a low score (ie, 1-3 level) for the clarity of the example PTO task and who preferred reimbursing treatment for neither of the patient groups. We identified only 11 "speeders", and hence did not perform sensitivity analyses on this subgroup.

Before conducting the analyses, we formulated the following hypotheses:

Hypothesis 1: Respondents prefer—and attach a higher weight to—reimbursing treatment for the relatively more severely ill patient group, when patients' ages are equal (module 1).
Hypothesis 2: Respondents prefer—and attach a higher weight to—reimbursing treatment for the relatively younger patient group, when patients' severity levels are equal (module 2).

As evidence regarding the direct trade-off between severity and age is limited,[17,26] we formulated no hypothesis, yet explored preferences when patients' severity level and age varied simultaneously (module 3).

We conducted the analyses using Stata 15.1 (Stata Corp LP, College Station, TX).

Results

Respondents assessed the clarity of the concepts in text at mean (SD) 5.9 (1.0) and in graphs at 5.7 (1.2), and of the example choice task at 5.5 (1.3) and example PTO task at 5.0 (1.7) on the 7-point Likert scale. Of the respondents, 177 (17.3%) reported a low score for the clarity of the example PTO task.

Table 3 presents the percentages of respondents with a preference for reimbursing treatment for group A, group B, or neither of the groups. Table 4 presents the PTO ratios.

In module 1, most respondents (range 47.4%-58.9% across choice sets; Table 3) preferred and, on average, attached a higher weight to reimbursing treatment for the relatively more severely ill patient group at all ages. These results provide evidence in support of hypothesis 1. The strength of respondents' preferences varied across severity levels, ages, and the size of the difference in severity between the groups. When the difference in severity was "small" (choice sets 1 to 6; Table 4), the MOR was 2.51, indicating that respondents, on average, valued reimbursing treatment for 1 more severely ill patient equally to that for 2.51 less severely ill patients. The ROMs were relatively higher when the 2 patient groups were lower on the severity scale. The ROMs increased with age when patient groups were lower on the severity scale (choice sets 2, 4, and 6; Table 4). When the difference in severity was "large" (choice sets 7 to 9; Table 4), the MOR was 2.75 and the ROM was somewhat higher in the youngest age group.

In module 2, most respondents (range 44.4%-63.3% across choice sets; Table 3) preferred and, on average, attached a higher weight to reimbursing treatment for the relatively younger patient group at each severity level. These results provide evidence in support of hypothesis 2. The strength of preferences again varied across ages, severity levels, and the size of the difference in age between the groups. When the difference in age was "small" (choice sets 1 to 6; Table 4), the MOR was 2.46, indicating that respondents, on average, valued reimbursing treatment for 1 younger patient equally to that for 2.46 older patients. The ROMs were relatively higher when it concerned comparisons between the higher age groups (choice sets 2, 4, and 6; Table 4). The ROMs did not show a clear relation to severity levels. When the difference was "large" (choice sets 7 to 9; Table 4), the MOR was 3.50 and the ROM was slightly higher for the highest severity level.

In module 3, most respondents (range 38.5%-74.0% across choice sets; Table 3) preferred reimbursing treatment for the relatively younger patient group, and respondents, on average, attached a higher weight to reimbursing treatment for younger patients, irrespective of patients' severity levels (choice sets 1, 3, 5, and 7 have ROMs >1, whereas choice sets 2, 4, and 8 have ROMs <1). In module 3, the MORs were somewhat lower than in modules 1 and 2. When the difference in severity and age was "small" (choice sets 1 to 8;

Table 5. OLS regression results.*

DV: ΔnrPatients*QOLgain	Model 1†				Model 2‡			
	Including outliers		Excluding outliers		Including outliers		Excluding outliers	
	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
ΔSeverity	5245.40§	(1982.73, 8508.07)	1193.73¶	(556.33, 1831.14)	2215.53§	(663.32, 3767.73)	985.00¶	(450.69, 1519.31)
ΔAge	−2772.48#	(−6035.15, 490.19)	−968.02§	(−1605.42, −330.61)	−2051.20§	(−3603.41, −498.99)	−1490.30¶	(−2024.61, −956.00)
R^2	0.01		0.02		0.01		0.04	
N	1014		1005		1014		1006	

ΔAge indicates difference in age (in years) between patient group A and B; ΔnrPatients*QOLgain, difference in the number of patients between patient group B and A, each multiplied by the treatment-related gain in QOL; ΔSeverity, difference in severity (in points of QOL loss) between patient group A and B; CI, confidence interval; DV, dependent variable; OLS, ordinary least squares; QOL, health-related quality of life; QOLgain, treatment-related gain in QOL (20 points on the 0-100 visual analogue scale); SE, standard error.
*Respondents who completed the person-trade-off tasks in less than 115 seconds (n = 11) are excluded from this analysis. The results are presented including and excluding respondents with outlying (−1.96 > z-score > 1.96) preferences.
†Model 1 presents the results for a difference of 30 points in QOL loss and 30 years in age between the patient groups.
‡Model 2 presents the results for a difference of 60 points in QOL loss and 60 years in age between the patient groups.
§$P < .01$.
¶$P < .001$.
#$P < .10$.

Table 4), the MOR was 1.98, and when this was "large" (choice sets 9 to 10; Table 4), the MOR was 2.42.

Table 5 presents the OLS results. The ΔSeverity coefficient indicates that, ceteris paribus, respondents attached a higher weight to reimbursing treatment for relatively more severely ill patients. The implied ratios are 2.48 and 1.97 per point increase in ΔSeverity, given a "small" and "large" difference in severity between the groups in models 1 and 2, respectively. The ΔAge coefficient indicates that, ceteris paribus, respondents attached a higher weight to reimbursing treatment for relatively younger patients. The implied ratios are 1.49 and 1.73 per year increase in ΔAge, given a "small" and "large" difference in age between the groups in models 1 and 2, respectively. The MRS (95% CI) between ΔSeverity and ΔAge were −1.23 (−4.02, 1.56) and −0.66 (−1.01, −0.31) for models 1 and 2. This indicates that ΔAge was considered relatively more important when the differences between the patient groups were large. The sensitivity analyses showed consistent results.

Our results suggest that AS, PS, and PH are all consistent with societal preferences when severity levels differ between patient groups and their ages are equal. However, none of these severity operationalizations are fully consistent with societal preferences when severity levels are equal and ages differ between patient groups. When both severity levels and ages vary, our results suggest that AS may be most consistent with societal preferences (in 6/10 choice sets), followed by PS (in 2/10 choice sets), and PH (in 1/10 choice sets).

Discussion

In this study, we examined the relative strength of societal preferences for severity and age in the context of healthcare priority setting. Our main findings are that respondents, on average, preferred—and attached a higher weight to—reimbursing treatment for relatively more severely ill and younger patients, when preferences for severity and age were elicited separately. When preferences were elicited simultaneously, respondents, on average, preferred—and attached a higher weight to—reimbursing treatment for relatively younger patients, irrespective of patients' severity levels. We found that the relationship between preferences for severity and age is nonlinear and dependent on severity levels, ages, and the size of the difference in severity and/or age between the patient groups. Moreover, we found that preferences were generally stronger when the difference between the groups was larger. It should be noted that these findings need to be considered in relation to the applied design and that preferences for severity and age may differ when elicited on full (or equally wide) severity and age scales or when elicited in relation to prolonged health improvements or improved life expectancy. It should also be noted that a considerable minority of respondents stated a preference for reimbursing treatment for neither of the patient groups or for the relatively less severely ill or older patient group. This preference heterogeneity is consistent with the results of previous studies.[5]

Our results suggest that AS, PS, and PH are all consistent with societal preferences when severity levels differ between patient groups and their ages are equal, yet none are fully consistent with societal preferences when severity levels are equal and ages differ between patient groups. When both severity levels and ages vary, our results suggest that AS may be most consistent with societal preferences, followed by PS and PH. These latter results are consistent with those of Stolk et al,[26] who compared support for AS, PS, and PH by applying a ranking exercise in a small convenience sample (n = 65). Our results are also consistent with those of other studies that empirically investigated equity weights and emphasise the relevance of both patients' severity and age in allocation decisions (see, eg, refs [24,26–33,35,36]). For example, our results are consistent with those of Dolan and Tsuchiya,[28] who found that preferences for age were stronger than those for severity when preferences for both were elicited simultaneously.[28] They already suggested that these preferences, therefore, should not be considered in isolation of each other. Our results are inconsistent with some studies that found the social value of QALYs to be relatively independent of patients' age.[37-39] Although these differences may result from differences in applied preference-elicitation methods, they may also result from preference heterogeneity over time and between countries.

The main strengths of this study lie in the extensively pilot-tested questionnaire and innovative application of the PTO approach in a design in which we varied severity levels and ages separately and simultaneously between patient groups. Other methods (eg, a discrete choice experiment) could have been applied here; however, we chose to apply the PTO approach as this enabled us to keep the tasks relatively simple for respondents and to better understand the elicited preferences at the level of an individual respondent. Other strengths lie in the reduction of the possible risk of order effects by presenting 2 modules, the choice sets, and patient groups in random order and in the avoidance of other considerations by limiting the duration of the disease and QOL gain to 1 year. We are aware that this comes with the limitation that our results cannot be generalized to severity- and age-related preferences in relation to QALY gains. Some further limitations need to be discussed. A first limitation is common to all preference-elicitation studies and concerns the sensitivity of the results to the way in which the questions were framed.[5,28,40] For example, severity and age were the only attributes that varied, and this may have influenced the relative strength of preferences for these attributes. Preferences for severity and age may differ when elicited in combination with other attributes.[41] To some extent, this was already observed when age and severity were varied simultaneously rather than separately. A second limitation concerns the respondent instruction to assume that there were no other differences between the groups than those explicated. Nonetheless, preferences may have been influenced by omitted variables.[41,42] For example, respondents may have considered aspects such as spillover effects on family or informal carers and productivity losses.[30,33] They may also have considered the acceptability of a less than perfect health state at an older age, the degree of health inequality between the groups, and the risk or uncertainty associated with severity levels and treatment gains.[43-46] A third limitation concerns a possible upward bias of the PTO ratios caused by censoring the number of patients between 100 and 1 000 000 in the PTO tasks.[47] However, the influence of censoring may be counterbalanced by including respondents with no preference for either of the groups in the analysis as the robustness checks indicated that the obtained ratios were in fact pulled downward by this. A final limitation concerns the low R^2 values of the regression models. These low values indicate that the models do not explain much of the data variability. Nevertheless, the implied ratios associated with an increase of 30 points and 30 years in the difference in severity and age between the patient groups resemble the mean-based PTO ratios in Table 4 and the OLS results aided the interpretation of the PTO ratios in module 3 as intended.

Our results are consistent with those of previous studies, suggesting that the outcomes of allocation decisions informed by AS may be better aligned with societal preferences than those informed by PS or PH.[20,26,27-31] However, our results also suggest that none of these severity operationalizations may be sufficiently adequate for guiding decisions that concern patients of different ages. At least, not when aiming to align the outcomes of these decisions with societal preferences. If so, current AS, PS, and PH applications may all need to be age weighted when ages differ between patient groups, irrespective of differences in their severity level. This may be even more necessary for those severity operationalizations that do not indirectly favor younger over older patients, like AS does. Whether or not age weighting is considered normatively acceptable and societal preferences are to be reflected in allocation decisions are pressing questions in this context, but fall outside the scope of the current article.

Further research is warranted to examine the relative strength of preferences for severity and age when severity is operationalized on a QALY scale, thus incorporating both length and QOL attributes. Research may further be aimed at examining the relevant thresholds for QALYs gained in different age-adjusted severity classes. For this, a matrix-based design could potentially be used to account for the nonlinear relationship between preferences for severity and age.[16] Finally, different preference-elicitation methods and the inclusion of other attributes in the trade-off may lead to different findings. Hence, further research is necessary to examine the most appropriate design for estimating severity- and age-based equity weights for informing allocation decisions in healthcare.

Conclusions

The results of this study indicate that the public prefers prioritizing relatively more severely ill patients when patients' ages are equal and younger patients when their severity levels are equal. When both patients' severity levels and ages vary, the public seems to prefer prioritizing relatively younger patients, irrespective of patients' severity levels. These results suggest that, when aiming to better reflect societal preferences in decision making, current severity operationalizations and decision-making frameworks may need to be adjusted for age-related societal preferences.

Acknowledgments

The authors thank Lucas Goossens for his assistance with the regression analyses. The usual disclaimer applies.

Source of financial support

This study is part of a larger project examining the broader societal benefits of healthcare and was funded by a consortium of Pfizer, GlaxoSmithKline, AbbVie, Amgen, and AstraZeneca in the Netherlands. The funders had no role in the study design, data collection and analysis, interpretation of the data, preparation of the article, or decision to submit for publication. The views expressed in this article are those of the authors.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2019.07.012.

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 principal of proportional shortfall. *Eur J Health Econ*. 2013;14(1):107–115.
2. Ottersen T, Mæstad O, Norheim OF. Lifetime QALY prioritarianism in priority setting: quantification of the inherent trade-off. *Cost Eff Resour Alloc*. 2014;12(1):2.
3. Hernæs UJ, Johansson KA, Ottersen T, Norheim OF. Distribution-weighted cost-effectiveness analysis using lifetime health loss. *Pharmacoeconomics*. 2017;35(9):965–974.
4. Sabik LM, Lie RK. Priority setting in health care: lessons from the experiences of eight countries. *Int J Equity Health*. 2008;7(1):4.
5. Shah K. Severity of illness and priority setting in healthcare: a review of the literature. *Health Policy*. 2009;93(2-3):77–84.
6. Schwappach DL. Resource allocation, social values and the QALY: a review of the debate and empirical evidence. *Health Expect*. 2002;5(3):210–222.
7. Nord E. The person-trade-off approach to valuing health care programs. *Med Decis Making*. 1995;15:201–208.
8. Williams A. Intergenerational equity: an exploration of the 'fair innings' argument. *Health Econ*. 1997;6(2):117–132.
9. 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.
10. 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.
11. Norwegian Ministry of Health and Care Services. *Principles for Priority Setting in Health Care: Summary of a White Paper on Priority Setting in the Norwegian Health Care Sector*. Oslo, Norway: Norwegian Ministry of Health and Care Services; 2017.
12. legemiddelverk Statens. Beregning av absolutt prognosetap I hurtig metodevurderinger av legemidler–bruke beregningsprinsippet for behandling eller forebygging? Notat fra en tverretatlig arbeidsgruppe. https://legemiddelverket.no/Documents/Offentlig%20finansiering%20og%20pris/Dokumentasjon%20til%20metodevurdering/APT-%20behandling%20eller%20forebygging_310119.pdf. Accessed February 17, 2019.

13. Magnussen J, Aaserud M, Granaas T, et al. På ramme alvor—Alvorlighet og prioritering. https://www.regjeringen.no/contentassets/d5da48ca5d1a4b128c72fc5daa3b4fd8/paa_ramme_alvor.pdf. Accessed June 2, 2019.
14. Magnussen J. Severity of illness and priority setting in Norway: summary of a report from a working group. https://www.regjeringen.no/contentassets/d5da48ca5d1a4b128c72fc5daa3b4fd8/summary_the_magnussen_report_on_severity.pdf. Accessed June 2, 2019.
15. Nederland Zorginstituut. *Ziektelast in de praktijk. Diemen,*. The Netherlands: Zorginstituut Nederland; 2018.
16. 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.
17. Stolk EA, Van Donselaar G, Brouwer WBF, Van Busschbach JJV. Reconciliation of economic concerns and health policy. *Pharmacoeconomics*. 2004;22(17):1097–1107.
18. Stolk EA, Goes E, Kok E, Van Busschbach JJV. Uitwerking criteria noodzakelijkheid, eigen rekening en verantwoording en lifestyle. College voor Zorgverzekeringen. Breedte geneesmiddelenpakket, 01/54. Amstelveen. The Netherlands: College voor Zorgverzekeringen; 2001.
19. Versteegh MM, Corro Ramos C, Buyukkaramikli A, et al. Severity adjusted probability of being cost-effective. *Pharmacoeconomics*. 2019;37(9):1155–1163.
20. Zorginstituut Nederland. *Kosteneffectiviteit in de praktijk. Diemen*. The Netherlands: Zorginstituut Nederland; 2015.
21. Dolan P, Olsen JA. Equity in health: the importance of different health streams. *J Health Econ*. 2001;20(5):823–834.
22. Brazier J, Rowen D, Mukuria C, et al. Eliciting societal preferences for burden of illness, therapeutic improvement and end of life for value based pricing: a report of the main survey. http://eprints.whiterose.ac.uk/99493/1/Eliciting%20societal%20preferences%20-%20011.pdf. Accessed July 7, 2018; 2013.
23. 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):281–288.
24. Nord E. Concerns for the worse off: fair innings versus severity. *Soc Sci Med*. 2005;60(2):257–263.
25. Edlin R, Round J, McCabe C, et al. *Cost-Effectiveness Analysis and Ageism: A Review of the Theoretical Literature*. Leeds, UK: Leeds Institute of Health Sciences; 2008.
26. Stolk EA, Pickee SJ, Ament AHJA, Van Busschbach JJV. Equity in health care prioritisation: an empirical inquiry into social value. *Health Policy*. 2005;74(3):343–355.
27. 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.
28. 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.
29. Olsen JA. Priority preferences: "end of life" does not matter, but total life does. *Value Health*. 2013;16(6):1063–1066.
30. Bobinac A, Van de Wetering EJ, Van Exel NJA, Brouwer WBF. Equity-dependent social willingness to pay for a QALY. https://repub.eur.nl/pub/94502. Accessed October 7, 2018; 2015.
31. 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 Health Econ*. 2018;19(6):797–806.
32. Reckers-Droog VT, Van Exel NJA, Brouwer WBF. Who should receive treatment? An empirical enquiry into the relationship between societal views and preferences concerning healthcare priority setting. *PLoS One*. 2018;13(6):e0198761.
33. Bobinac A, Van Exel NJA, Rutten FF, Brouwer WBF. Inquiry into the relationship between equity weights and the value of the QALY. *Value Health*. 2012;15(8):1119–1126.
34. Oemar M, Janssen B. *EQ-5D-5L User Guide: Basic Information on How to Use the EQ-5D-5L Instrument*. Rotterdam, The Netherlands: EuroQol Group; 2013.
35. Johannesson M, Johansson PO. Is the valuation of a QALY gained independent of age? Some empirical evidence. *J Health Econ*. 1997;16:589–599.
36. Tsuchiya A, Dolan P, Shaw R. Measuring people's preferences regarding ageism in health: some methodological issues and some fresh evidence. *Soc Sci Med*. 2003;57(4):687–696.
37. Johansson-Stenman O, Martinsson P. 2008. Are some lives more valuable? An ethical preferences approach. *J Health Econ*. 2008;27:739–752.
38. Lancsar E, Wildman J, Donaldson C, Ryan M, Baker R. Deriving distributional weights for QALYs through discrete choice experiments. *J Health Econ*. 2011;30(2):466–478.
39. Van de Wetering EJ, Van Exel NJA, Bobinac A, Brouwer WBF. Valuing QALYs in relation to equity considerations using a discrete choice experiment. *Pharmacoeconomics*. 2015;33(12):1289–1300.
40. Ubel PA. How stable are people's preferences for giving priority to severely ill patients? *Soc Sci Med*. 1999;49(7):895903.
41. Van de Wetering L, Van Exel NJA, Rose J, Hoefman R, Brouwer WBF. Are some QALYs more equal than others? *Eur J Health Econ*. 2016;17(2):117–127.
42. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science*. 1981;211(4481):453–458.
43. Dolan P. The measurement of individual utility and social welfare. *J Health Econ*. 1998;17:39–52.
44. Brouwer WBF, Van Exel NJA, Stolk EA. Acceptability of less than perfect health states. *Soc Sci Med*. 2005;60(2):237–246.
45. Doctor JN, Miyamoto JM. Person tradeoffs and the problem of risk. *Expert Rev Pharmacoecon Outcomes Res*. 2005;5(6):677–682.
46. 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):79.
47. Ternent L, Tsuchiya A. A note on the expected biases in conventional iterative health state valuation protocols. *Med Decis Making*. 2013;33(4):544–546.