giving weight to equity: improving priority setting in healthcare

Vivian Reckers-Droog

Giving Weight to Equity: Improving priority setting in healthcare

Vivian Reckers-Droog

ISBN: 978-94-6416-455-8

Cover by Martien van den Hoek Art & Illustrations
Layout by John Droog
Printed by Ridderprint, the Netherlands

Giving Weight to Equity: Improving priority setting in healthcare

Gewicht geven aan rechtvaardigheid: Verbeteren van prioritering in de gezondheidszorg

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. F.A. van der Duijn Schouten

en volgens besluit van het College van Promoties.
De openbare verdediging zal plaatsvinden op

vrijdag 21 mei 2021 om 10:30 uur

door

Viviana Tamara Reckers-Droog
geboren te Utrecht

Erasmus University Rotterdam

Promotiecommissie:

Promotoren:

Prof.dr. W.B.F. Brouwer

Prof.dr. N.J.A. van Exel

Overige leden:

Prof.dr. D.M.J. Delnoij

Prof.dr. J.J. Polder

Dr. G.A. de Wit

Voor mijn vader en moeder,
aan wie ik alles te danken heb.

Table of Contents

Chapter 1
Introduction

1.1 Introduction

The demand for healthcare is rapidly increasing for reasons that include ageing populations, the availability of increasingly advanced and expensive new health technologies, and higher standards of living that raise the expectations of health and healthcare [1–3]. The growth rate of expenditures in healthcare tends to exceed that of the economy [4,5], which jeopardises the sustainability of publicly financed healthcare systems and risks the crowding out of other collective expenditures, for example, on public order and safety and on education [6]. As healthcare resources are limited, the resulting pressure on the available budgets renders priority setting in the allocation of healthcare resources inevitable. The need for priority setting is widely recognised and explicitly addressing priority setting is necessary for an optimal allocation of healthcare resources. Nevertheless, the debate and controversy that often follow (in particular, negative) resource-allocation decisions illustrate that healthcare priority setting is still politically and societally sensitive [1,2,7].

Publicly financed healthcare systems have two important objectives [1,8]. The first objective is to generate as much (health) value as possible from the healthcare budget, and hence to allocate the available resources in an efficient manner. The second objective is to distribute health and healthcare fairly, and hence to allocate the available resources in an equitable manner. Although it has been argued that an optimal allocation of healthcare resources involves setting priorities that contribute to meeting both objectives [1,8], important questions remain about which equity considerations should be considered, what (relative) weight these considerations should receive, and how they should be incorporated in resource-allocation decisions.

This thesis aims to contribute to a better understanding of societal concerns for equity in healthcare priority setting, in particular for priority setting based on disease severity and the age of patients. The background to this thesis is described in sections 1.2 to 1.5 and the research questions it addresses are outlined in section 1.6. Note that this thesis has a strong focus on healthcare priority setting in the Netherlands. Nonetheless, its findings also have relevance for other countries that seek to integrate societal concerns for equity with concerns for efficiency into the decision-making process.

1.2 Economic evaluations of new health technologies

Economic evaluations of new health technologies are increasingly used to inform decision makers on how to allocate the available healthcare resources in an efficient manner. In economic evaluations, health gains from health technologies are often expressed in terms of quality-adjusted life-years (QALYs). QALYs capture treatment-related gains in both quality of life and life expectancy and combine these gains into a single outcome measure [9]. The quality-of-life component of the QALY is measured on an interval scale,

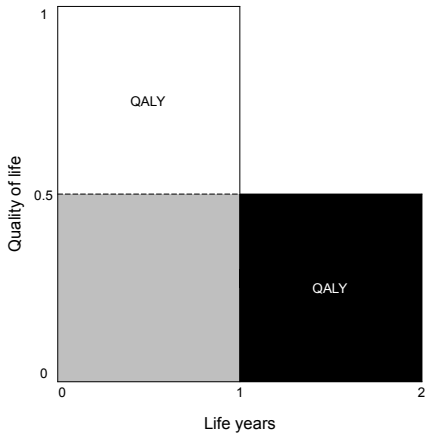


Fig. 1.1 Quality-adjusted life-year (QALY)

on which the utility associated with the health state 'dead' is by convention anchored at 0 and the utility associated with the health state 'full health' at 1 [9]. Health states that are perceived as being worse than 'dead' are associated with disutility (i.e. quality of life <0). Figure 1.1 illustrates that one QALY can represent one year in full health or, for example, two years in a less good health state with a quality of life of 0.5.

In economic evaluations, QALY gains and the relevant costs associated with generating these gains are compared between the new health technology and a reference case (e.g. current standard of care) [9]. The outcome of this comparison is the incremental cost-effectiveness ratio (ICER) of the new health technology, which is used to inform resource-allocation decisions in healthcare [9].

Economic evaluations of new health technologies are commonly conducted from a broad societal perspective or a more narrow healthcare perspective [2,9]. In countries that apply a societal perspective (e.g. the Netherlands), the underlying objective is to maximise social welfare from the healthcare budget [2]. Therefore, the broader impacts of a resource-allocation decision (i.e. the benefits and costs that fall outside the healthcare system) are taken into account in the decision-making process [2]. In these countries, the decision rule can be written as [2,10]:

$$v_Q \cdot \Delta Q - \Delta ct > 0 \qquad \text{(Eq. 1.1)}$$

where v_Q denotes the monetary value of a QALY (i.e. its consumption value), ΔQ the incremental QALY gain, and Δct the incremental total costs that are associated with the QALY gain [2,10]. Note that ct is the sum of healthcare costs (ch) and broader consumption costs (cc) [2]. This equation can be rewrit-

ten to demonstrate the decision rule that is commonly applied in economic evaluations:

$$\Delta ct/\Delta Q < v_Q \quad \text{(Eq. 1.2)}$$

This equation demonstrates that the allocation of resources towards a new health technology can be considered welfare improving when the ICER of a new health technology (i.e. Δct/ΔQ) is lower than the threshold value specified in terms of the societal willingness to pay for a QALY [2,10].

In countries that apply a healthcare perspective (e.g. England), the objective is to maximise population health from a (fixed) healthcare budget [2,10]. Therefore, only the impact of a resource-allocation decision on the healthcare system is taken into account in the decision-making process. In these countries, ct is replaced by ch and v is replaced by k in the decision rules, where k denotes the monetary value of a QALY specified in terms of the opportunity costs of resource-allocation decisions in healthcare [2,10–12].

The classic approach in economic evaluations is to adhere to the principle that a "QALY is a QALY is a QALY" [13], meaning that all QALY gains are valued equally regardless of by whom and in which context QALYs are gained [14]. However, adhering to this principle has become the subject of much debate as it relies on the assumption of distributive neutrality, whilst accumulating evidence suggests that the social value of a QALY may vary depending on equity considerations associated with characteristics of the patients (e.g. their age, lifestyle, and socioeconomic status), disease (e.g. its prevalence, severity, and outcome), and health technology under evaluation (e.g. the type and size of health gains it generates) [14–17]. In response to this debate, it has been advocated to incorporate equity considerations into the decision-making framework [14–16,18,19].

Many countries have incorporated equity considerations into the decision-making framework, albeit often in an ad hoc, implicit manner [20]. For example, by accepting a higher ICER in case a new health technology is indicated for severely ill patients or by requiring a lower co-payment from severely ill patients to improve their access to a new health technology [21,22]. However, to facilitate transparent and consistent decision-making [20], an increasing number of countries incorporates such considerations in an explicit manner by applying equity weights in economic evaluations [23–25]. One of the first countries to apply such weights was the Netherlands.

1.3 Equity weighting in economic evaluations

Equity weights can be attached to QALY gains or be reflected in the monetary threshold value (v or k) used in economic evaluations [1,14,16,26–29]. In the former case, the equity-adjusted ICER of a health technology is evaluated against a fixed monetary threshold value. In the latter case, the ICER of a

health technology is evaluated against a flexible, equity-adjusted monetary threshold value. The equity-adjusted decision rule can be written as [1,29]:

$$\Delta ct/\Delta Q_i < v_{Qi} \quad \text{(Eq. 1.3)}$$

where the subscript i denotes the equity characteristic of the incremental QALY gain and v_{Qi} the monetary threshold value specified in terms of the societal willingness to pay for the equity-adjusted QALY [1,29]. Considering that the equity characteristic i has direct consequences for the distribution of health and healthcare, its normative justification and the empirical support for the underlying (combination of) equity consideration(s) are highly important. Furthermore, the equity-adjusted decision rule (Eq. 1.3) highlights the need to define a base case 'equity scenario' of which the weight and monetary value are known to enable differentiation between QALY gains [1,26,30]. This also raises questions about the implications of applying different weights or values for the allocation of healthcare resources.

The equity weights that are currently applied in economic evaluations are all based on some definition of disease severity that can be derived from the renowned severity and fair innings approaches [23–25,31–33]. These approaches have the same normative standpoint that a higher weight should be attached to health gains in those who are worse off in terms of health in order to reduce health inequalities in society [31–33]. However, they have different standpoints with regard to *whom* are considered worse off. According to the severity approach, those with a lower amount of prospective health are worse off and, therefore, should be given a higher weight in resource-allocation decisions [32]. However, according to the fair innings approach, those with a lower amount of lifetime (i.e. past *and* prospective) health are worse off and should be given a higher weight in such decisions [31,33]. Although both approaches are normatively justifiable and to some extent empirically supported, evidence suggests that neither approach is fully aligned with societal concerns for equity weighting based on the disease severity of patients [1,34]. Indeed, evidence suggests that the public considers it important to take patients' prospective *as well as* their lifetime health into account in resource-allocation decisions [1,15,34–36].

In an attempt to balance societal concerns for the severity and fair innings approaches, the intermediate approach 'proportional shortfall' was introduced and gradually implemented into the decision-making framework in the Netherlands [25,34,37].

1.4 Proportional shortfall

Proportional shortfall is calculated as the fraction of patients' disease-related QALY loss, relative to their remaining QALY expectation in absence of the disease and is measured on a scale ranging from 0 "no QALY loss" to 1 "com-

plete loss of remaining QALYs" (i.e. immediate death) [34]. According to this approach, those who lose a larger fraction of their remaining QALY expectation (i.e. those with a higher level of proportional shortfall) are worse off in terms of health and should, therefore, be given a higher weight in resource-allocation decisions [25,34]. In the Netherlands, this is operationalised by evaluating the ICER of a new health technology against a flexible, equity-adjusted monetary threshold value that is positively associated with the proportional shortfall level of the patients for whom the technology is indicated [25,38].

An important reason for implementing proportional shortfall in the Netherlands was that it mitigated discrimination against older patients implied by the fair innings approach and the use of ICERs for informing resource-allocation decisions [25,37]. In theory, proportional shortfall indeed does not discriminate on the basis of age between patients [34]. For example, in the case of immediate death, patients aged 10 and 80 will both have a proportional shortfall level of 1 and are, therefore, given the same (high) weight in resource-allocation decisions. However, in decision-making practice, proportional shortfall may not just mitigate discrimination against older patients, but it may pave the way for discrimination in favour of older patients as they are, ceteris paribus, more likely to lose a larger fraction of their remaining QALY expectation than younger patients. For example, when patients aged 10 and 80 both lose two of their remaining QALYs, patients aged 80 will have a relatively higher level of proportional shortfall and, therefore, are given a higher weight than patients aged 10 in resource-allocation decisions.

Equity weights based on proportional shortfall (or on any other definition of disease severity that can be derived from the severity and fair innings approaches) do not explicitly distinguish between patients of different ages, nor aim to give weight to patients' age in resource-allocation decisions. Nevertheless, the weights may be inextricably related to patients' age, and hence their application in healthcare priority setting may have different consequences for patients of different ages. This raises the question to what extent this approach aligns with societal concerns for equity weighting based on disease severity *and* the age of patients.

Proportional shortfall combines aspects of the severity and fair innings approaches and is, therefore, related to established conceptions of equity in healthcare priority setting. It should however be noted that there is still much debate about what equity approach is considered best for informing resource-allocation decisions in healthcare. Likewise, there are still questions about the empirical support in the general public for the use of specific approaches, such as proportional shortfall. The latter will be addressed in this thesis.

1.5 Societal concerns for equity in healthcare priority setting

There are different ways to obtain insight into societal concerns for equity in healthcare priority setting and to incorporate related preferences into the decision-making framework. Insight into societal preferences is often obtained by using preference-elicitation methods (e.g. person trade-off and contingent-valuation tasks) in purposively designed questionnaires that are administered to large, representative samples of the general public [39–41]. The elicited preferences can, for example, be incorporated into the decision-making process by using them as an empirical base for equity weighting in economic evaluations.

In order to increase the legitimacy and accountability of resource-allocation decisions as well as public support for potentially unpopular decisions, insight into societal preferences is increasingly obtained by using deliberative methods (e.g. citizens panels and juries), whether or not in combination with equity weighting in economic evaluations [40,42–45]. Deliberative methods facilitate a two-way flow of information between decision makers and (a small sample of) the public and serve to transform the viewpoints and preferences of both parties by acts of dialogue and negotiation [40,46,47]. As such, these methods can help to ensure that not only the outcomes of resource-allocation decisions are normatively justifiable and empirically supported, but that the decision-making process also meets these requirements [42,45].

Despite the increased use of deliberative methods in healthcare priority setting, their impact has rarely been assessed empirically [40]. Hence, questions remain about whether and how deliberative methods influence the viewpoints and preferences of participants and to what extent they (continue to) represent those of the general public [40,41,48]. Insight into the effect of deliberative methods is indispensable for making informed decisions on whether, how, and at what stage of the decision-process deliberative methods are best incorporated into the decision-making framework.

1.6 Objective and outline of this thesis

The overall objective of this thesis is to contribute to the improvement of the decision-making framework by providing further insight into societal concerns for equity in healthcare priority setting, in particular for priority setting based on disease severity and the age of patients.

To meet the overall objective of this thesis, the following research questions are addressed:

1. What is the normative justification and empirical support for equity weighting based on proportional shortfall in the Netherlands?

2. How much weight does the public attach to disease severity and the age of patients in healthcare priority setting?

3. What is the public willing to pay for quality-of-life gains in patients with different ages and levels of disease severity?

4. What is the public willing to pay for quality-of-life and life-expectancy gains in patients with different ages at the end of life?

5. How do different viewpoints on healthcare priority setting relate to concerns for equity and efficiency in resource-allocation decisions?

6. How does participating in a deliberative citizens panel influence the viewpoints of participants on healthcare priority setting?

Chapter 2 examines the normative justification and empirical support for equity weighting based on proportional shortfall in the Netherlands. A key finding in this chapter is that empirical support for equity weighting based on proportional shortfall may be inextricably related to the age of the patients for whom a new health technology is indicated. Chapters 3 to 5 build on this finding and examine how much weight the public attaches to disease severity and the age of patients in healthcare priority setting. Chapter 3 presents the results of a study in which person trade-off tasks were applied in an innovative manner to obtain severity- and age-based equity weights. Chapters 4 and 5 present the results of studies in which contingent-valuation tasks were applied to examine what the public is willing to pay for health gains in patients with different ages and levels of disease severity, the latter also operationalised in an end-of-life context. Chapter 6 examines how different viewpoints on healthcare priority setting relate to concerns for equity and efficiency in healthcare priority setting. Chapter 7 examines how participating in a deliberative citizens panel influences the viewpoints of participants on healthcare priority setting by extending previous applications of Q methodology.

Chapters 2 to 7 are based on articles published (or submitted for publication) in international peer-reviewed journals and can therefore be read independently. It should be noted that this implies some inevitable overlap between these chapters.

Chapter 8 is the final chapter of this thesis. It discusses the main findings of this thesis, its strengths and limitations, and the implications for policy and future research.

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 Health Econ. 2013;14(1):107–15.
2. Brouwer W, van Baal P, van Exel N, Versteegh M. When is it too expensive? Cost-effectiveness thresholds and health care decision-making. Eur J Health Econ. 2019;20(2):175–80.
3. Rosén P, Karlberg I. Opinions of Swedish citizens, health-care politicians, administrators and doctors on rationing and health-care financing. Health Expect. 2002;5(2):148–55.
4. Stadhouders N, Koolman X, van Dijk C, Jeurissen P, Adang E. The marginal benefits of healthcare spending in the Netherlands: Estimating cost-effectiveness thresholds using a translog production function. Health Econ. 2019;28(11):1331–44.
5. Pammolli F, Riccaboni M, Magazzini L. The sustainability of European health care systems: Beyond income and aging. Eur J Health Econ. 2012;13(5):623–34.
6. Hicks LL. Making hard choices. J Leg Med. 2011;32(1):27–50.
7. Sabik LM, Lie RK. Priority setting in health care: Lessons from the experiences of eight countries. Int J Equity Health. 2008;7:1–13.
8. Maynard A. European health policy challenges. Health Econ. 2005;14(S1):255–63.
9. 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.
10. Gravelle H, Brouwer WBF, Niessen L, Postma M, Rutten FFH. Discounting in economic evaluations: Stepping forward towards optimal decision rules. Health Econ. 2007;(16):307–17.
11. 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.
12. 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.
13. Weinstein MC. A QALY is a QALY is a QALY - Or is it? J Health Econ. 1988;7(3):289–90.
14. Wailoo A, Tsuchiya A, McCabe C. Weighting Must Wait. Pharmacoeconomics. 2009;27(12):983–9.
15. Schwappach DL. Resource allocation, social values and the QALY: A review of the debate and empirical evidence. Health Expect. 2002;5(3):210–22.
16. Cookson R, Drummond M, Weatherly H. Explicit incorporation of equity considerations into economic evaluation of public health interventions. Health Econ Policy Law. 2009;4(2):231–45.
17. 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;126:128–37.

18. 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.
19. Shah KK, Tsuchiya A, Wailoo AJ. Valuing health at the end of life: An empirical study of public preferences. Eur J Health Econ. 2014;15(4):389–99.
20. Chalkidou K, Glassman A, Marten R, Vega J, Teerawattananon Y, Tritasavit N, et al. Priority-setting for achieving universal health coverage. Bull World Health Organ. 2016;94(6):462–7.
21. Le Polain M, Franken M, Koopmanschap M, Cleemput I. Drug reimbursement systems: international comparison and policy recommendations. Health Services Research (HSR). Brussels: Belgian Health Care Knowledge Centre (KCE); 2010.
22. Broqvist M, Sandman L, Garpenby P, Krevers B. The meaning of severity – do citizenś views correspond to a severity framework based on ethical principles for priority setting? Health Policy. 2018;122(6):630–7.
23. National Institute for Health and Care Excellence. Guide to the Methods of Technology Appraisal. London; 2013.
24. Norwegian Ministry of Health and Care Services. Principles for priority setting in health care. Oslo; 2017.
25. Zorginstituut Nederland. Ziektelast in de praktijk. Diemen; 2018.
26. 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):157–71.
27. 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):1119–26.
28. Round J, Paulden M. Incorporating equity in economic evaluations: a multi-attribute equity state approach. Eur J Health Econ. 2018;19(4):489–98.
29. Versteegh MM, Ramos IC, Buyukkaramikli NC, Ansaripour A, Reckers-Droog VT, Brouwer WBF. Severity-Adjusted Probability of Being Cost Effective. Pharmacoeconomics. 2019;37(9):1155–63.
30. Wagstaff A. QALYs and the equity-efficiency trade-off. J Health Econ. 1991;10(1):21–41.
31. Williams A. Intergenerational equity: An exploration of the "fair innings" argument. Health Econ. 1997;6(2):117–32.
32. Nord E. Concerns for the worse off: Fair innings versus severity. Soc Sci Med. 2005;60(2):257–63.
33. Harris J. The Value of Life: Introduction to Medical Ethics. London: Routledge & Keegan Paul; 1985.
34. 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):1097–107.
35. 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.
36. Shah KK. Severity of illness and priority setting in healthcare: A review of the literature. Health Policy. 2009;93(2–3):77–84.

37. Stolk EA, Goes E, Kok E, Van Busschbach JJV. Uitwerking criteria noodzakelijkheid, eigen rekening en verantwoording en lifestyle. In: Breedte geneesmiddelenpakket. Amstelveen: College voor Zorgverzekeringen; 2001.

38. Zorginstituut Nederland. Cost-effectiveness in practice. Diemen; 2015.

39. Ryan M, Scott DA, Reeves C, Russell EM, Napper M. Eliciting Public Preferences for Healthcare: A Systematic Review of Techniques. Chinese J Evidence-Based Med. 2009;9(2):129–30.

40. Mitton C, Smith N, Peacock S, Evoy B, Abelson J. Public participation in health care priority setting: A scoping review. Health Policy. 2009;91(3):219–28.

41. Dolan P, Cookson R, Ferguson B. General practice of priority setting in health care: focus group study. BMJ. 1999;318:916–9.

42. Daniels. Accountability For Reasonableness: Establishing A Fair Process For Priority Setting Is Easier Than Agreeing On Principles. Acad Manag Rev. 2000;31(2):386–408.

43. Daniels N, Sabin JE. Accountability for reasonableness: An update. BMJ. 2008;337:a1850.

44. Abelson J, Forest P-G, Eyles J, Smith P, Martin E, Gauvin F-P. Obtaining public input for Past experiences and future prospects. Can Public Adm. 2002;45(1):70–97.

45. Abelson J, Forest PG, Eyles J, Smith P, Martin E, Gauvin FP. Deliberations about deliberative methods: Issues in the design and evaluation of public participation processes. Soc Sci Med. 2003;57(2):239–51.

46. Rowe G, Frewer LJ. A typology of public engagement mechanisms. Sci Technol Human Values. 2005;30(2):251–90.

47. Abelson J. Understanding the role of contextual influences on local healthcare decision making: Case study results from Ontario, Canada. Soc Sci Med. 2001;53(6):777–93.

48. Abelson J, Eyles J, McLeod CB, Collins P, McMullan C, Forest PG. Does deliberation make a difference? Results from a citizens panel study of health goals priority setting. Health Policy. 2003;66(1):95–106.

Chapter 2
Looking back and moving forward: On the application of proportional shortfall in healthcare priority setting in the Netherlands

Based on: 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.

Abstract

The increasing demand for healthcare and the resulting pressure on available budgets render priority setting inevitable. If societies aim to improve health and distribute health(care) fairly, equity-efficiency trade-offs are necessary. In the Netherlands, proportional shortfall (PS) was introduced to quantify necessity of care, allowing a direct equity-efficiency trade-off. This study describes the history and application of PS in the Netherlands and examines the theoretical and empirical support for PS as well as its current role in healthcare decision making. We reviewed the international literature on PS from 2001 onwards, along with publicly accessible meeting reports from the Dutch appraisal committee, Adviescommissie Pakket (ACP), from 2013 to 2016. Our results indicate that there is support for the decision model in which necessity is quantified and incremental cost-effectiveness ratios are evaluated against associated monetary reference values. The model enables a uniform framework for priority setting across all healthcare sectors. Although consensus about the application of PS has not yet been reached and alternative ways to quantify necessity were found in ACP reports, PS has increasingly been applied in decision making since 2015. However, empirical support for PS is limited and it may insufficiently reflect societal preferences regarding age and reducing lifetime-health inequalities. Hence, further investigation into refining PS—or exploration of another approach—appears warranted for operationalising the equity-efficiency trade-off.

2.1 Introduction

The demand for healthcare is rapidly increasing for reasons that include ageing populations and the availability of increasingly advanced and expensive (new) health technologies. As healthcare resources remain scarce, the resulting pressure on available budgets renders healthcare priority setting inevitable [1,2]. Although politically and societally sensitive, the need for prioritisation is widely recognised and explicitly addressing priority setting has become indispensable for developing fairer methods for resource allocation in healthcare [3,4].

Economic evaluations of health technologies are often used to inform decision makers regarding how to allocate healthcare resources in an optimal way for society. However, the outcomes of economic evaluations only predict such decisions to a moderate extent [3,5,6]. One explanation for this disparity is that decision makers are not exclusively concerned with maximising health given available budgets, but also with distributing health(care) equitably and fairly [3,5,7,8]. Hence, an optimal allocation of resources involves setting priorities that contribute to both efficiency and equity in the distribution of health(care) [9]. Recognising that these are both important objectives of healthcare systems, it has been advocated that societal concerns for equity be explicitly and transparently incorporated into the decision-making framework [10–12].

In economic evaluations, the value of a health technology is commonly expressed in terms of an incremental cost per quality-adjusted life-year (QALY) ratio (ICER) that is evaluated against some monetary threshold value per QALY gained [3,13–15]. When the ICER is below this threshold, a health technology is considered cost-effective and eligible for reimbursement [16]. The classic approach in the economic-evaluation framework is to value QALY gains equally, i.e. to adhere to the principle that a ”QALY is a QALY is a QALY”, regardless of beneficiary and health technology characteristics [17]. However, this approach has been highly debated as it relies on the assumption of distributive neutrality [3]. In response to this debate, two general approaches have been suggested for operationalising the equity-efficiency trade-off [3,5]. One of these approaches applies equity weights to QALY gains and evaluates the adjusted ICER against a fixed monetary threshold value, and the other evaluates an unadjusted ICER against a flexible monetary threshold value [3,5,16]. Ideally, the operationalisation of the equity-efficiency trade-off is both normatively justifiable and empirically supported. However, this proves to be neither easy nor straightforward [3,18].

In relation to the operationalisation of the equity-efficiency trade-off, the severity of illness (SOI) and fair innings (FI) equity approaches have attracted much attention internationally. According to the normative theories about distributive justice that underlie these approaches, priority should be given to

those who are worse off in terms of health [11,19,20]. However, the approaches are based on different normative arguments with regard to whom is considered worse off, and hence differ with regard to how they are operationalised [3,19]. A common operationalisation of SOI aims to equalise absolute health benefits in terms of current and prospective health, while FI aims to do so in terms of lifetime health [3,10,11,19]. As such, FI also considers past health [11,19]. Although both SOI and FI are to some extent normatively justifiable and empirically supported, neither of these approaches appears to satisfactorily reflect societal preferences for equity [3,5,7]. Nonetheless, different countries have either implicitly or explicitly developed normative principles or guidelines that include (aspects of) SOI or FI for informing allocation decisions in healthcare [3,5,7]. For example, in the United Kingdom (UK), the National Institute for Health and Care Excellence (NICE) formalised the SOI approach by launching guidelines for prioritising end-of-life care [21,22], and in Norway, the SOI approach is currently formalised in terms of absolute shortfall [23,24]. In an attempt to balance societal concerns regarding SOI and FI [5], proportional shortfall (PS) was introduced in the Netherlands as an equity approach that combines aspects of SOI and FI [3,5]. Although consensus about the application of PS has not yet been reached [25], the approach received considerable support from politicians and policy makers and was incorporated into the assessment phase of healthcare priority setting in the Netherlands [1,3]. As such, the Netherlands is one of the first countries to explicate the equity criterion in this context [3,5].

This study describes the history and application of PS in the Netherlands and examines the theoretical and empirical support for PS as well as its current role in healthcare decision making in the Netherlands by reviewing the international PS literature and publicly accessible meeting reports from the Dutch appraisal committee, the Adviescommissie Pakket (ACP). Although this study primarily focuses on healthcare priority setting in the Netherlands, the results of the study may also be useful for other countries seeking to operationalise the equity-efficiency trade-off for informing allocation decisions in healthcare.

2.2 A brief history of healthcare priority setting in the Netherlands

The report "Choices in health care" that was presented by the Dunning Committee in 1991 was a landmark publication on healthcare priority setting in the Netherlands [26]. In this report, four criteria for priority setting were formulated: necessity, effectiveness and efficiency of care, and patients' individual responsibility for (paying for) care. In this report, the Dunning Committee used the metaphor of a funnel to describe a criteria-based decision model for evaluating the composition of the publicly funded health-insurance package. Based on this hierarchical model, technologies that (would subsequently) pass all criteria were to be included in the basic benefits package. The report was pivotal for the discussion on priority setting, and in the following years, the

criteria were put into practice [2,6,27]. The Dutch National Health Care Institute (ZIN) later reformulated these criteria as necessity of care, effectiveness, cost-effectiveness, and necessity of insurance, respectively, and supplemented these with a feasibility criterion [1,28]].

Although none of the criteria were defined and operationalised without dispute, this proved to be particularly difficult for the necessity of care criterion [2,29]. The Dunning Committee defined necessity of care as care that is necessary for the prevention of premature death and/or for patients who—due to some disease or condition—cannot function normally in society [2,26]. The latter part of this definition was regarded as problematic, as it was unclear how to interpret and quantify 'normal' functioning. Moreover, the term 'necessity' implied an absolute rather than a relative cut-off point for decision making, which was amplified by the Dunning Committee's use of a funnel metaphor [2,26]. If a technology failed to pass 'the sieve of necessity', the technology would not be incorporated into the public health-insurance package, and assessment of its (cost-) effectiveness and need for insurance would be superfluous [2]. However, as the degree to which health technologies are necessary varies, it was suggested that this criterion be regarded as neither absolute nor isolated from the other criteria [2,30,31].

In 2001, Stolk et al. [2] proposed a decision model in which necessity of care was defined as 'burden of illness' (BOI) and operationalised as a relative criterion by attaching a higher necessity score to health technologies that target diseases with a higher BOI level. Stolk et al. [2] described BOI as the average disease-related loss in quality and length of life of patients, relative to the situation in which the disease had been absent and quantified BOI in terms of QALYs on a 0−1 scale. Furthermore, they proposed connecting the necessity of care and (cost-) effectiveness criteria by attaching a higher societal willingness to pay (WTP) per QALY gained to a higher level of BOI. Specifically, the authors suggested dividing the continuous 0−1 BOI scale into seven categories and evaluating the ICER of (new) health technologies against seven associated monetary threshold values per QALY gained. The proposed cost-effectiveness threshold values per QALY gained ranged from approximately €4,500 to €45,000 [32]. Deciding on the exact cut-off points for the BOI categories, the cost-effectiveness threshold range, and the shape of their reciprocal relationship were regarded as matters of political and societal concern.

The proposed model received broad support as it contributed to the development of a transparent and coherent decision model for healthcare priority setting in the Netherlands by explicitly connecting the criteria formulated by the Dunning Committee and enabling a uniform and systematic quantification of BOI across patient groups and disease areas [2,3,27–29]. Between 2002 and 2005, BOI was further formalised as proportional shortfall (PS) and defined as a principle that is based on the normative standpoint that priority in healthcare should be given to those who, due to some disease and if left

untreated, lose the largest proportion of their QALY expectancy in absence of the disease [5,32,33]. PS is measured on a scale from 0 (no QALY loss) to 1 (complete loss of remaining QALY, i.e. immediate death), by applying:

$$PS = \frac{\text{Disease-related QALY loss}}{\text{Remaining QALY expectation in absence of the disease}} \quad \text{(Eq. 2.1)}$$

For example, a disease that results in the loss of 30 out of 60 remaining QALYs has a PS level of 0.5 (30/60), and a disease that results in the loss of 60 out of 80 remaining QALYs has a PS level of 0.75 (60/80). The remaining QALY expectation in absence of the disease can be calculated from age- and sex-specific mortality data [2,5]. Equation 1 can be rewritten as:

$$PS = 1 - \frac{\text{Expected QALYs without treatment}}{\text{Remaining QALY expectation in absence of the disease}} \quad \text{(Eq. 2.2)}$$

Applying Equation 2.2 to the previous example, the PS level of 0.5 is calculated as 1 - (30/60), and the PS level of 0.75 is calculated as 1 - (20/80). PS can also be calculated by using the number of expected QALYs 'with current treatment' rather than 'without treatment' in the equations [34]. This may be a more logical calculation of PS as it arguably uses a more relevant comparator and hence agrees with the economic-evaluation methodology. However, it should be noted that calculating PS relative to the current treatment will likely lead to a different, specifically lower, PS level for the same beneficiaries and (new) health technologies. Consequently, the outcome of a reimbursement decision that is informed by a PS level that is calculated relative to the current treatment may be different for the same beneficiaries and (new) health technologies than when the decision is informed by a PS level that is calculated relative to having no treatment. The debate on the preferred comparator is likely to continue in the coming period.

While consensus concerning the definition and operationalisation of BOI gradually increased, its exact categories and the associated cost-effectiveness threshold range remained a subject of discussion for some time. In 2006, the Council for Public Health and Society (RVZ) suggested a continuous, upward-sloping curve with a maximum reimbursement of €80,000/QALY [29]. This figure was substantiated by the World Health Organisation (WHO) rule of thumb that less than three times the GDP per capita per disability-adjusted life-year (DALY) averted indicated good value for money for a health technology [35], by the finding that most reimbursed health technologies in the UK had an ICER of approximately €79,000/QALY [36], and by estimations of the value of a statistical life [37,38]. Although the figure of €80,000/QALY may have been set somewhat arbitrarily, it was considered "reasonable" [29,39]. Moreover, even though €80,000/QALY was not officially adopted as the threshold value at that time, it was influential and provided the basis for ZIN to set three BOI categories with a maximum reimbursement of €80,000/QALY

Table 2.1 Maximum reference values (in €) per QALY gained [25]

Burden of illness	Maximum reference value per QALY gained
0.10 – 0.40	€ 20,000
0.41 – 0.70	€ 50,000
0.71 – 1.00	€ 80,000

for the highest BOI category in 2015 [25]. Table 2.1 presents these three BOI categories and the associated monetary reference values and shows that a higher WTP per QALY gained, i.e. a higher necessity score, is attached to health technologies that target diseases with a higher BOI level [2,25]. For example, the ICER of a health technology that targets a disease with a PS level of 0.5 is evaluated against a reference value of €50,000/QALY, while the ICER of a health technology that targets a disease with a PS level of 0.75 is evaluated against a reference value of €80,000/QALY. A health technology that targets a disease with a PS level below 0.1 is, in principle, not considered for reimbursement. Hence, this category is not included in the table [25,29]. Whether it is feasible, in practice, to not reimburse a health technology that targets a disease with a low PS level remains doubtful [40]. For example, episodic illnesses like migraine may not lead to a high average PS, but do represent substantial shortfall during the episode.

Given the maximum reimbursement of €80,000/QALY for the highest BOI category and the intention to associate increasing levels of BOI with increasing monetary reference values, ZIN set the two lower thresholds at €20,000 and €50,000 per QALY. Together these may be seen as forming a logical set of values, given the endpoint of €80,000/QALY in relation to the highest BOI. In relation to the other two values, ZIN also referred to the threshold value that is applied to national immunisation and preventive care programmes in the Netherlands (€20,000/QALY threshold) and to a Dutch study on the societal WTP per QALY gained 'in others' (€50,000/QALY) [25,35]. ZIN advised reassessing the reference values every five to ten years [25] and to not use them as strict cut-off values, but rather as references for the Dutch government when conducting price negotiations with pharmaceutical companies and for the ACP when recommending incorporation of health technologies into the public health-insurance package.

The model, in which BOI is quantified and the ICERs of health technologies are evaluated against associated reference values, enables a transparent and coherent decision-making framework. Given that this model is increasingly applied in the Netherlands, the question arises whether there actually is sufficient support for the operationalisation of BOI in terms of PS to explicate the equity criterion. In the next two sections we will discuss the theoretical and empirical support for using PS to inform priority setting in healthcare. In the subsequent section we will review the current role of PS in healthcare decision making in the Netherlands.

2.3 Is there theoretical support for proportional shortfall?

In order to optimally allocate healthcare resources for society, it has been advocated that societal concerns regarding equity be incorporated in the decision-making framework [10–12]. However, what society considers to be equitable and fair for priority setting is a normative question that different people in different contexts may answer differently. Hence, when operationalising the equity-efficiency trade-off, an additional trade-off between different societal concerns regarding equity must be made. Consequently, increasing equality in the distribution of health(care) by applying one equity approach may lead to increasing inequality in the context of applying another [3,41]. It has also been argued that, when operationalising the equity–efficiency trade-off, different operationalisations are bound to face corresponding difficulties [3,16]. For example, in the context of curative healthcare, questions may arise concerning the handling of episodic diseases and the quantification of related health benefits [3,16]. In the context of preventive healthcare, questions may arise concerning the group of beneficiaries and the timeframe that is regarded as relevant for estimating health benefits [3,16,18]. For example, should PS be calculated for all treated persons or only those for whom the illness was prevented? And should PS be calculated from the time of the preventive treatment or from the time the prevented illness would otherwise have occurred? Such choices can have a profound effect on the outcomes of PS calculations [42]. Other questions may, for example, arise concerning the use of age- and sex-specific mortality data as a reference point or threshold for calculating PS [20,43,44]. The use of such different reference points for different (age and sex) groups implies that there is not one age or health expectancy that would serve as a normative reference level for all groups. Hence, this could be regarded as including some inequ(al)ities in the calculation of PS [20]. These issues illustrate that not only is the choice of an equity approach normative, but additional normative choices must be made when applying the chosen equity approach in practice [3]. Inevitably, these choices have a large impact on PS calculations and therefore may have distributional consequences [3,18]. Although some initial choices were made when operationalising PS in the Netherlands [1,2,26], it should be noted that the discussion about how best to solve these issues is ongoing (both in the context of healthcare priority setting in the Netherlands and internationally).

SOI and FI are two renowned equity approaches that are based on different normative arguments regarding whom is considered worse off in terms of health [3,19]. As described earlier, SOI commonly aims to equalise health benefits in terms of current and prospective health, and FI aims to do so in terms of lifetime health [3,10,11,19]. As such, FI is consistent with the notion that, all else equal, younger people should be prioritised over older people as they have not yet enjoyed a fair share of lifetime health [5,11]. It should be noted that the role of age is merely indirect in the FI approach as it is applied

as a proxy for lifetime health [20,43,44]. Indeed, in the FI approach, age itself is not regarded as a morally relevant argument for priority setting [44].

PS does not originate from a unique theory about distributive justice but was developed as an equity approach that combines aspects of SOI and FI by prioritising those who are worse off in terms of a lower amount of prospective and lifetime health [3,5]. While SOI and FI aim to equalise absolute health benefits, PS aims to equalise relative benefits between persons with respect to their potential for health [5,33]. It has been argued that PS balances societal concerns regarding SOI and FI and treats the two approaches as equally important [5]. However, PS is calculated as the fraction of disease-related QALY loss relative to the remaining QALY expectation in absence of the disease rather than to the lifetime-QALY expectation from birth. Various authors have discussed the relative nature of PS and the theoretical and empirical relevance of using a lifetime perspective for informing allocation decisions in healthcare [20,23,45–47]. Here, we would like to point out that PS may be viewed as placing more emphasis on relative prospective-health loss, i.e. the SOI component of PS, than on relative lifetime-health loss, i.e. the FI component of PS. This is illustrated by the fact that PS does not, by definition, discriminate between people with different levels of 'enjoyed' lifetime health as healthcare beneficiaries of all ages could potentially experience the same level of PS. For example, in the case of immediate death, healthcare beneficiaries who are 10 and 80 years old are given the same weight in the distribution of healthcare, as both will have a PS level of 1. However, when the same beneficiaries lose two of their remaining QALYs, more weight will be given to the 80-year olds, as their PS level will be higher than that of the 10 year olds. Indeed, in allocation decisions, PS may more frequently give a higher weight to older patients than the FI approach would. Stolk et al. [5] argued that the FI approach "discriminates against the elderly more strongly than policy makers seem to prefer" and that PS could mitigate the ageism that is implied by the FI approach. It was, therefore, hypothesised that PS might be better aligned with distributional preferences of health policy makers. Should this hypothesis not be supported by empirical evidence, the authors suggested to add age weights and adjust PS for age-related preferences.

A strength of PS, which it shares with the SOI and FI approaches, lies in its quantification of health losses in terms of QALYs. This enables the application of PS across disease areas and patient populations. However, this strength comes with a limitation as treatment benefits beyond health and health-related quality of life (QOL) that may not be captured by the QALY are increasingly recognised as being relevant [48]. Therefore, the current application of PS, i.e. its quantification in terms of QALYs, may be regarded as appropriate for informing decisions concerning curative and preventive treatments but less so for decisions concerning treatments that focus on broader benefits, for example related to wellbeing [49]. If the aim is to generate social welfare from the public health-insurance package, the application of an equity approach

that is uniformly applicable and hence that models information concerning health, QOL, and broader wellbeing could be preferable for informing decisions concerning all healthcare sectors. We stress that this limitation should not be attributed to PS (or to FI or SOI) as a principle but rather to the way in which PS is currently quantified and applied in decision-making practice. In fact, PS does enable a uniform decision model for priority setting across all healthcare sectors, as the QALY in the PS equation can be replaced with—or complemented by—any other (generic) outcome measure of choice.

2.4 Is there empirical support for proportional shortfall?

We examined empirical support for PS by reviewing the international literature on PS in the context of healthcare priority setting. We used the search terms "proportional shortfall", "preference", "elicitation", "priority setting", and "health" or "healthcare" in Google Scholar. The search was performed on October 16, 2017 and supplemented with a hand search. We restricted the search to articles written in English or Dutch, published between 2001, i.e. the year in which PS was introduced in the Netherlands, and 2017, and of which the full text was available. Articles were selected for review if the aim of the study was to elicit preferences for PS relative to either preferences for no equity weighting or for weighting on the basis of another equity approach, such as SOI and/or FI. Our search resulted in 205 studies, in seven of which preferences for PS were elicited. Table 2.2 presents an overview of these seven studies and their results.

Stolk et al. [33] compared support for SOI, FI, and PS by asking respondents to assign a priority rank to the treatment of ten health conditions. Stolk et al. found strong evidence for PS being consistent with social preferences for healthcare priority setting. Although preferences for PS dominated preferences for SOI, stronger support was found for FI. The authors obtained these results using a small convenience sample in the Netherlands that consisted of health policy makers, researchers, and students. Consequently, the results may be prone to bias, e.g. due to respondents sharing common opinions.

Olsen [50] examined support for PS in a sample that was representative of the general adult population in Norway in terms of age and sex. Olsen applied a pairwise-choice task and asked respondents to prioritise patients based on their age, remaining lifetime health without treatment, and increase in remaining lifetime health with treatment. Olsen found strong support for the FI approach; however, he found no support for PS.

Brazier et al. [51] examined support for BOI operationalised in terms of PS in a sample that was representative of the general population in the UK in terms of age and sex by performing a web-based discrete choice experiment (DCE). Their main results did not support PS. However, when respondents who see-

mingly misunderstood the DCE task were excluded, some support for PS was found.

Rowen et al. [54] used the data from Brazier et al. [51] to examine support for PS by applying the number of expected QALYs 'with current treatment' rather than 'without treatment' in the PS equation. Rowen et al. concluded that, although the results were not robust against different versions of the DCE survey, there was some modest support for BOI operationalised in terms of PS relative to the current treatment.

Van de Wetering et al. [52] examined support for PS in a sample that was representative of the general adult population in the Netherlands in terms of age, sex, and education level by conducting a web-based DCE. They found substantial preference heterogeneity and some counterintuitive results, as respondents were less likely to prioritise patients with higher levels of PS.

Bobinac et al. [53] examined societal WTP for QALY gains in patients with different levels of PS in a sample that was representative of the general adult population in the Netherlands in terms of age, sex, and education level by conducting a web-based survey. They found occasional support for PS as a predictor of the WTP for QALY gains. Some support for PS was found when QALY gains were relatively small. However, the level of support decreased when QALY gains increased in size. In addition, support for PS was generally dominated by concerns for the (younger) age of patients.

Table 2.2 Empirical evidence on support for proportional shortfall (2001–2017)

Study	Year	Country	Design[a]	N	Sample	Support for PS[b]
Stolk et al. [33]	2005	NL	Ranking exercise	65	Convenience	++
Olsen [50][c]	2013	NO	Pairwise-choice task	503	General public (age and sex)	--
Brazier et al. [51]	2013	UK	DCE	3,669	General public (age and sex)	--/-
Van de Wetering et al. [52]	2015	NL	DCE	1,205	General public (age, sex, and education level)	--
Bobinac et al. [53]	2015	NL	WTP	1,320	General public (age, sex, and education level)	-
Rowen et al. [54][d]	2016	UK	DCE	3,669	General public (age and sex)	+
Richardson et al. [55][c]	2017	AU	CSPC	606	General public (age)	+

AU, Australia; CSPC, constant sum paired comparison; DCE, discrete choice experiment; NL, the Netherlands; NO, Norway; PS, proportional shortfall; UK, United Kingdom; WTP, willingness to pay; [a] Mode of administration: web-based survey in all studies; [b] Level of support for PS indicated by -- = no, - = limited, + = modest, ++ = strong; [b] Olsen [50] and Richardson et al. [55] examined support for PS in the context of preferences for length of life; [d] Rowen et al. [54] examined support for PS relative to the current treatment.

Richardson et al. [55] examined support for PS in a sample that was close to being representative of the general adult population in Australia in terms of age. They applied constant-sum paired comparison tasks and asked respondents to prioritise patients based on their gain in life years due to treatment, age, years to death without treatment, and age at death with and without treatment. Their study found some support for PS; however, found that concerns for PS were dominated by concerns for the (individual) SOI and FI approaches. Richardson et al. further found that PS insufficiently reflects respondents' age-related preferences.

Although each of these studies examined societal support by eliciting preferences for PS, it is important to note that the studies differ with respect to the samples, methods, additionally included variables, and/or equity approaches. In addition, Olsen [50] and Richardson et al. [55] examined preferences for equity in the context of length of life, and hence did not present PS in terms of proportional QALY shortfall. Consequently, a direct comparison of the results presented in Table 2.2 is not possible.

2.5 What is the role of proportional shortfall in healthcare decision making?

The necessity (of care and of insurance), effectiveness, cost-effectiveness, and feasibility criteria are addressed and quantified in the assessment phase of healthcare decision making in the Netherlands and subsequently assessed on social and ethical grounds in the appraisal phase. If the operationalisation of BOI in terms of PS is considered suboptimal for explicating the equity criterion, it seems reasonable to expect that this would be explicitly discussed during meetings of the ACP appraisal committee.

To examine the current role of PS in the appraisal phase of healthcare decision making in the Netherlands, we conducted a review of publicly accessible ACP meeting reports that were published between 1 January 2013 and 31 December 2016. The reports include agendas, minutes, and documents, including decision reports and draft ZIN reports that were discussed by the ACP. Table 2.3 presents the terms (and their domains) addressing healthcare priority setting that we used for searching the reports (in the Dutch language, but translated here for clarity). Reports that did not allow a digital search, including ACP reports that were published before 1 January 2013 were excluded from the review, as were search terms that occurred in the names of health organisations and government ministries. Draft versions of minutes were included only if final minutes were not published.

Between 2013 and 2016, 179 ACP reports were published of which two were excluded for not allowing a digital search. Table 2.4 presents the frequency with which the search terms were identified in the remaining 177 reports. The necessity of care and of insurance, effectiveness, cost-effectiveness (inclu-

Table 2.3 Search terms used for reviewing ACP meeting reports

Domain	Search term
Priority-setting criteria	Necessity of care Necessity of insurance Effectiveness; effect Cost-effectiveness; efficiency Feasibility
Equity considerations	Severity of illness Fair innings Burden of illness Absolute shortfall Proportional shortfall
Treatment benefits	Therapeutic outcome; therapeutic value (Health-related) quality of life Quality-adjusted life-year; QALY Wellbeing Capability Life satisfaction
Patient characteristics	Age Socio-economic status; SES Lifestyle Culpability; individual responsibility
Reference values	Reference value(s) (Monetary) threshold

ACP, Adviescommissie Pakket (the healthcare appraisal committee in the Netherlands)

ding the search term efficiency), and feasibility criteria were identified 1,680, 495, 8,700, 4,423, and 236 times, respectively. The effectiveness criterion was most frequently found, followed by the cost-effectiveness and necessity of care criteria. The necessity of insurance and feasibility criteria were identified less frequently.

PS was identified 14 times in a total of six reports, four of which discussed the operationalisation of BOI in terms of PS. In a report from 2013, ZIN described the definition and calculation of PS. In this report, ZIN stated that "PS was developed at a time when ageism was an important issue in the allocation of healthcare resources" and that "therefore BOI is calculated in proportion to life expectation, which ensures that PS does not distinguish between younger and older people". However, ZIN also stated that "recently, there are increasing indications that people do discriminate between age groups" and that people "value health gains in younger people more than in older people", which "argues against PS and the rule of rescue, and in favour of FI". In the accompanying minutes, an ACP member stated that "the passage about BOI is still not in agreement with what was discussed in previous meetings" and that s/he "understand[s] that applying the capability approach is out of reach", but that s/he "would like to see the denominator removed from the presented definition of PS". In a report from 2015, ZIN stated that "because we have not

Table 2.4 Frequency of terms addressing healthcare priority setting in ACP meeting reports published between 2013 and 2016

Year	Type	*n*	Necessity of care	Necessity of insurance	Effectiveness	Cost effectiveness	Feasibility	SOI	FI	BOI	AS	PS	QOL	Well being	Age	SES	Lifestyle	Reference values
2013	Agenda	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Minutes[a]	9	10	2	53	28	1	0	0	6	0	1	10	1	2	1	2	0
	Documents[b]	35	357	121	1820	572	20	0	8	191	0	5	339	28	323	22	137	6
2014	Agenda	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Minutes	7	6	2	50	46	1	0	0	0	0	0	6	0	0	1	1	0
	Documents[b]	31	378	74	3626	1428	67	0	0	91	0	0	846	30	246	14	183	31
2015	Agenda	8	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
	Minutes	8	10	4	104	84	1	0	1	11	0	0	46	4	9	0	1	12
	Documents[b]	34	239	109	1781	1673	72	0	2	237	0	7	725	12	119	6	23	194
2016	Agenda	6	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Minutes	6	18	4	103	73	2	0	0	17	0	0	35	3	6	0	2	19
	Documents[b]	17	659	179	1163	517	71	0	0	119	0	1	381	5	190	10	37	66
Frequency			1680	495	8700	4423	236	0	11	672	0	14	2388	83	895	54	386	328
Number of reports		177	112	64	121	120	55	0	4	68	0	6	89	28	79	17	43	26
Mean frequency per report			15.0	7.7	71.9	36.9	4.3	0	2.8	9.9	0	2.3	26.8	3.0	11.3	3.2	9.0	12.6

ACP, Adviescommissie Pakket (the healthcare appraisal committee in the Netherlands); AS, absolute shortfall; BOI, burden of illness; FI, fair innings; PS, proportional shortfall; QOL, health-related quality of life; SES, socio-economic status; SOI, severity of illness; Cost-effectiveness is the pooled result of "cost-effectiveness" and "efficiency"; Effectiveness is the pooled result of "effect", "effectiveness", and "therapeutic outcome/value"; Lifestyle is the pooled result of "lifestyle", "individual responsibility", and "culpability"; QOL is the pooled result of "(health-related) quality of life" ,"quality-adjusted life-year", and "QALY"; Reference values is the pooled result of "reference value(s)" and "(monetary) threshold; Wellbeing is the pooled result of "wellbeing", "capability", and "life satisfaction"; [a] Minutes for 2013 include one concept version; [b] Documents include meeting and decision reports.

yet reached consensus about the quantification of BOI, we will temporarily […] quantify BOI in terms of DALYs". In this report, ZIN additionally stated that "priority will be given to solving this issue" and that "a report on the quantification of BOI will be issued this summer". In a later report from 2015, ZIN stated that "the next coming months will be used to see how to better align the equity criterion PS with current social preferences". According to the accompanying minutes, these statements by ZIN were not discussed by ACP members.

The reference of one of the ACP members to Sen's capability approach [41] may indicate a preference for quantifying health benefits in terms of broader wellbeing, as for example is done by applying the ICECAP measure [56]. Wellbeing, including the terms capability and life satisfaction, was identified 93 times in 29 reports from 2013 onwards, among which the capability approach was identified 15 times in five reports (not in table). In these reports, the capability approach was discussed as an alternative to quantifying health benefits in terms of QALYs. In a report from 2013, ZIN stated that "a recent discussion involves the question of whether the capability approach is better aligned with the social basis that underlies managing the public health-insurance package" and that "applying this approach may be more appropriate for healthcare sectors where 'health gains' are not the primary objective, such as long-term care and mental healthcare". The same report stated that "changing the desired outcome of healthcare does not answer the question of when care is more necessary for one person than for another" and that "the capabilities approach can also be applied to calculate lifetime capabilities (fair innings), prospective capabilities, or the relative loss of capabilities (proportional shortfall)", and so "applying the capability approach will not solve the issue of prioritisation in healthcare".

The ACP member's request to remove the denominator from the PS equation may indicate a preference for operationalising the equity criterion in terms of absolute shortfall (AS) rather than proportional shortfall, and this may in turn indicate a preference regarding FI, age, and reducing lifetime-health inequalities [24,50]. AS was not identified in any of the ACP reports and the FI approach was identified 11 times in four reports. In contrast, the SOI approach was identified 0 times. However, concerns for SOI that were expressed through concerns for prospective-health loss, the rule of rescue, and for the severity of (symptoms of) a disease or condition were identified 3 times in 1 report, 5 times in 3 reports, and 2614 in 92 reports, respectively (not in table). Age was identified 895 times in 79 reports. Regarding age and other patient characteristics, age was identified 16.6 and 2.3 times more frequently than SES and lifestyle (including the search terms culpability and individual responsibility), respectively.

Although the operationalisation of BOI in terms of PS was occasionally discussed in some reports, and in one report from 2014 an ACP member stated

that "BOI cannot be quantified in numeric terms in this specific situation", the application of BOI itself was not discussed. BOI was identified 672 times in 68 reports, and from 2015 onwards, increasingly related to the corresponding reference values (see Table 4.1), which were identified 328 times in 26 reports. In 2013, BOI was most frequently expressed in qualitative terms, e.g. in terms of "low" or "high" BOI, and only sometimes in numeric terms, by disability weights used for calculating DALYs. From 2014 onwards, BOI was less frequently expressed in qualitative terms and was mostly quantified by disability weights or the number of DALYs lost, which at times were presented alongside the mean life expectancy of patients with and without the disease. In three reports, BOI was addressed as a relative measure; however, the presented disability weights or DALYs lost were not applied as such. From 2015 onwards, BOI was most frequently quantified in terms of PS, in a total of seven reports. In four of these reports, PS calculations were presented alongside disability weights, number of DALYs lost, mean life expectancy, or years of life lost calculations.

Based on these results, it appears that the application of BOI was not publicly discussed by ACP members between 2013 and 2016. The operationalisation of BOI in terms of PS, and the role of PS in healthcare decision making, was infrequently discussed. While BOI was most frequently expressed qualitatively in 2013, it was increasingly quantified in later years, usually in terms of disability weights or the number of DALYs lost. From 2015 onwards, ICERs were increasingly related to the monetary reference values per QALY gained that were set by ZIN that year [25] and BOI was most frequently quantified in terms of PS. In this context, it needs noting that there was a change in ACP members in 2015 and this may have contributed to the increased application of PS from then on. In the reports, PS calculations were frequently presented alongside disability weights, number of DALY lost, life expectancy, and years of life loss calculations. This variety may reflect that there is no consensus (yet) about the application of PS in healthcare decision making in the Netherlands.

2.6 Discussion

The importance of operationalising the equity-efficiency trade-off for informing priority setting in healthcare is increasingly recognised. This study described the history and application of PS in the Netherlands, examined the theoretical and empirical support for PS as an operationalisation of the equity-efficiency trade-off, and looked into the current role of PS in healthcare decision making.

Overall, our results indicate general support for the decision model in which BOI is quantified and the ICERs of health technologies are evaluated against the reference values per QALY gained set by ZIN. This model connects the criteria for healthcare decision making that were previously formulated by the Dunning Committee and enables a uniform decision model for priority setting

across all healthcare sectors. Consequently, the model has received broad support in research and policy circles and has been incorporated into the healthcare decision-making framework in the Netherlands.

Although the model is increasingly applied in decision-making practice, the results of our literature review suggest that theoretical support for PS is moderate at best. In applying PS, a trade-off between the SOI and FI approaches is made and, consequently, societal preferences for either of the two equity approaches may be insufficiently reflected when allocating resources in healthcare. However, this may be regarded as a general limitation that comes with applying any equity approach in practice, as improving equality in the distribution of health(care) by applying one equity approach may inevitably be associated with increasing inequality in the context of applying another [3]. A limitation that is not restricted, but may be more specific to applying PS, is that it mitigates ageism between patient groups, as beneficiaries of all ages can experience the same level of PS. The results of our study suggest that this may inadequately reflect societal preferences relating to age and reducing lifetime-health inequalities between patient groups. Although the results of our literature review suggest that empirical support for PS is limited, it should be noted that that empirical evidence regarding PS so far is scarce and inconclusive, so that a rejection of the PS approach can also not be concluded. The societal concern regarding age that is currently insufficiently reflected by PS could be incorporated by adjusting PS for age. However, there is no empirical evidence (to date) to support the hypothesis that this would better align with societal preferences and hence future research on this topic will be necessary. The results of our review of publicly accessible ACP reports suggest that the ACP did not publicly discuss the definition and operationalisation of necessity of care in terms of BOI between 2013 and 2016. In fact, BOI became increasingly quantified, and ICERs were increasingly evaluated against the reference values per QALY gained set by ZIN in 2015. The operationalisation of BOI in terms of PS was publicly discussed by the ACP, although only on rare occasions. This may indicate that the ACP supports the operationalisation of BOI in terms of PS. However, the variety of BOI quantifications in ACP reports demonstrates that consensus about the operationalisation and quantification of BOI has not yet been reached.

Relatively few studies have examined the theoretical and empirical support for PS and, to our knowledge, no other study has examined the current role of PS in healthcare decision making in the Netherlands. Although this limits our ability to compare our results with those of others, we would like to compare the results of our empirical literature review to the results of a study conducted by Nord and Johansen [57] and the public consultation of NICE on the valued-based assessment of health technologies [58–60]. Nord and Johansen [57] examined support for PS, relative to preferences for no other equity approach, by conducting an empirical literature review that built on an earlier review by Shah [61] and included 20 preference studies that were conducted

in nine different countries between 1991 and 2011. Under the assumption of a stable health condition and no loss in length of life for patients, Nord and Johansen found strong support for PS, although the strength of the support varied greatly between the included studies. Regarding the inclusion criteria for our literature review, five of the studies that Nord and Johansen included were conducted after the introduction of PS in the Netherlands. Of these five studies, two quantified health benefits in terms of QOL and three in terms of QALYs. However, none of the three latter studies elicited preferences for PS and, as such, were not included in our literature review. Although the results of our review seem to be discordant with Nord and Johansen's results, and it is worth mentioning that in two of the five aforementioned studies the public was found to be less likely to prioritise patients with higher levels of SOI [45,62], a direct comparison of results is not possible for reasons that are previously described.

NICE conducted a public consultation in the UK in 2014 on the topic of value-based assessment of health technologies [58–60]. NICE asked the public, including patients, economists, academics, and members of the pharmaceutical industry ten related questions. One of the questions concerned the extent to which the public regarded PS as an appropriate approach for quantifying BOI [58]. NICE received responses from 121 individuals and organisations, but no general agreement emerged [59]. In summary, the public regarded PS as a measure that is feasible and suitable for calculating BOI in terms of health and QOL impact in cases where a disease affects older patients. However, as in the ACP, there were concerns about PS not being a suitable measure for capturing broader aspects of BOI that are not included in the QALY. In addition, there were concerns about PS assigning a lower weight to the BOI of younger patients than to older patients due to differences in the PS denominator, i.e. the remaining QALY expectation in absence of the disease. More generally, there were concerns about the strong reliance on the QALY in health technology assessment and in the calculation of BOI, resulting in a possible double counting of benefits when setting priorities in healthcare [59]. Because of the lack of agreement that emerged from the public consultation, NICE decided to not change their current health-technology appraisal and end-of-life guidelines [60].

Some limitations of our study must be mentioned. A first limitation concerns the lack of a systematic review of the studies that we used to examine the theoretical and empirical support for PS. However, as the number of studies examining PS is limited, we believe that our review was comprehensive and that our results were not influenced by the lack of a systematic search. A second limitation concerns the use of publicly accessible ACP reports to examine the current role of PS in healthcare decision making in the Netherlands. In addition to public meetings, the ACP held closed meetings between 2013 and 2016, and the role of PS may have been discussed more frequently in these. However, the reports of these closed meetings are not publicly accessible and

therefore could not be included in our review of ACP reports. Although the role of PS may have been discussed more frequently during closed meetings, and the inclusion of closed meeting reports might have changed the review results, it seems reasonable to expect that any discussion of PS in a closed meeting would have been reflected in a public meeting where the decision making actually took place. A final limitation concerns the risk of double counting search terms due to a possible overlap in ACP meeting reports. This risk was reduced by excluding concept versions of minutes unless a final version was not available; however, this distinction could not always be made for meeting documents. For example, documents concerning the reimbursement of a specific treatment may have been discussed at more than one ACP meeting. As a result, a higher relative importance may have been assigned to some of the search terms. However, as the main objective of our review of ACP reports was to examine the role of PS in healthcare decision making, and PS was infrequently identified, we believe that the influence of possible double counting on conclusions drawn from the review is limited. Concerning the frequency with which the search terms were identified in the ACP meeting reports, we would like to additionally point out that these frequencies should be considered in the broader context of priority setting in the Netherlands. This broader context determines the agenda and the priorities that are set in ACP meetings and hence influences the frequency with which the search terms were identified. Apart from these limitations, we consider it a strength of our study that we have examined support for PS at three different levels, i.e. at a theoretical, empirical, and decision-making level. To our knowledge, this is the first study to examine support for PS in such an extensive manner.

Ideally, the operationalisation of the equity-efficiency trade-off are normatively justifiable and empirically supported. The various normative choices that need to be made in this context indicate that a trade-off or a compromise between different societal concerns regarding equity and fairness needs to be made. Consequently, the 'perfect' explication of the equity criterion may not exist, and PS, like any other explication, will have its strengths and limitations. The results of our study indicate that the decision model in which increasing levels of BOI are quantified and ICERs are related to the associated monetary reference values per QALY gained is supported and increasingly applied in decision-making practice. The operationalisation of BOI in terms of PS enables a uniform decision model for priority setting across all healthcare sectors that can be applied by replacing or complementing the QALY in the PS equation with a broader, wellbeing-related, generic outcome measure such as the ICECAP [56]. The results of our study also indicate that PS insufficiently reflects societal preferences regarding age and reducing lifetime-health inequalities between patient groups. Future research is needed to develop and examine alternative versions of PS, such as a version of PS that is adjusted for wellbeing- and age-related preferences. These could be compared to the current operationalisation of PS, also in terms of alignment with general public preferences. There are different possibilities for combining PS and age

in a preference-elicitation study. For example, it may be interesting to elicit preferences for a PS version in which the denominator of the PS equation consists of patients' lifetime-QALY expectation. It may also be interesting to elicit preferences for combinations of PS and lifetime health (or age). For this, a matrix combining different age and PS classes could be used. To increase comparability between studies, we recommend using a more standardised approach to eliciting equity weights. Related to this, we would like to mention that the variety of the ways in which BOI is quantified in ACP reports may, understandably, reflect the present lack of consensus about the application of PS. However, this variety is in itself undesirable as it hampers the transparency and comparability of BOI quantifications for different beneficiaries and, subsequently, of the relevant cost-effectiveness threshold of (new) health technologies in the Netherlands. ZIN is expected to publish a report on the standardisation of the BOI quantification later this year to solve this issue.

2.7 Conclusions

The results of this study suggest that there is support for the decision model in which BOI is quantified and ICERs are evaluated against increasing monetary reference values. Although consensus regarding the application of PS has not yet been reached, BOI is increasingly quantified in terms of PS in decision-making practice. As any (generic) outcome measure can be included in the PS equation, PS enables a uniform decision model for priority setting across all healthcare sectors. Empirical support for PS appears to be limited, as PS may insufficiently reflect societal preferences regarding age and reducing lifetime-health inequalities. Hence, further investigation into the refinement of PS—or exploration of another approach—appears warranted for operationalising the equity-efficiency trade-off in healthcare priority setting.

References

1. College voor Zorgverzekeringen. Breedte geneesmiddelenpakket. Diemen; 2001.
2. Stolk EA, Goes E, Kok E, Van Busschbach JJV. Uitwerking criteria noodzakelijkheid, eigen rekening en verantwoording en lifestyle. In: Breedte geneesmiddelenpakket. Amstelveen: College voor Zorgverzekeringen; 2001.
3. 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 Health Econ. 2013;14(1):107–15.
4. Sabik LM, Lie RK. Priority setting in health care: Lessons from the experiences of eight countries. Int J Equity Health. 2008;7:1–13.
5. 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):1097–107.
6. Franken M, Nilsson F, Sandmann F, De Boer A, Koopmanschap M. Unravelling drug reimbursement outcomes: A comparative study of the role of pharmacoeconomic evidence in dutch and swedish reimbursement decision making. Pharmacoeconomics. 2013;31(9):781–97.
7. Nord E. Cost-value analysis in health care: Making sense out of QALYs. New York: Cambridge University Press; 1999.
8. Wouters S, van Exel J, Baker R, B.F. Brouwer W. Priority to End of Life Treatments? Views of the Public in the Netherlands. Value Health. 2017;20(1):107–17.
9. Maynard A. European health policy challenges. Health Econ. 2005;14(S1):255–63.
10. Nord E. The person-trade-off approach to valuing health care programs. Med Decis Mak. 1995;15(3):201–8.
11. Williams A. Intergenerational equity: An exploration of the "fair innings" argument. Health Econ. 1997;6(2):117–32.
12. 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.
13. Zorginstituut Nederland. Richtlijn voor het uitvoeren van economische evaluaties in de gezondheidzorg. Diemen; 2016.
14. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. London; 2013.
15. Canadian Agency for Drugs and Technologies in Health. Guidelines for the economic evaluation of health technologies: Canada (3rd edition). Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.
16. 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):1119–26.
17. Weinstein MC. A QALY is a QALY is a QALY - Or is it? J Health Econ. 1988;7(3):289–90.
18. Rappange DR, Brouwer WBF, Rutten FFH, Van Baal PHM. Lifestyle intervention: From cost savings to value for money. J Public Health. 2010;32(3):440–7.

19. Nord E. Concerns for the worse off: Fair innings versus severity. Soc Sci Med. 2005;60(2):257–63.
20. Ottersen T. Lifetime QALY prioritarianism in priority setting. J Med Ethics. 2013;39(3):175–80.
21. National Institute for Health and Care Excellence. Appraising life-extending, end of life treatments. Vol. 16. 2009. p. 5–8.
22. National Institute for Health and Care Excellence. End of life care for infants, children and young people with life-limiting conditions: Planning and management. London; 2016.
23. Norwegian Ministry of Health and Care Services. Principles for priority setting in health care. Oslo; 2017.
24. Ottersen T, Førde R, Kakad M, Kjellevold A, Melberg HO, Moen A, et al. A new proposal for priority setting in Norway: Open and fair. Health Policy. 2016;120(3):246–51.
25. Zorginstituut Nederland. Cost-effectiveness in practice. Diemen; 2015.
26. Dunning A. Kiezen en delen; Rapport van de commissie Keuzen in de Zorg. Den Haag: Ministerie Welzijn, Volksgezondheid en Cultuur; 1991.
27. College voor Zorgverzekeringen. Pakketbeheer in de praktijk 3. Diemen; 2013.
28. College voor Zorgverzekeringen. Pakketbeheer in de praktijk. Diemen; 2006.
29. Raad voor de Volksgezondheid en Zorg. Zinnige en Duurzame Zorg. Zoetermeer; 2006.
30. Regeringsbeleid WR voor het. Volksgezondheidszorg WRR-rapport 52. Den Haag: Sdu Uitgevers; 1997.
31. van Winzum C. Verdeling door verdunning; rapport van de commissie Commissie Criteria Geneesmiddelenkeuze. Geldermalsen: Stichting Gezondheidszorg en Publicaties; 1994.
32. College voor Zorgverzekeringen. Vervolgonderzoek breedte geneesmiddelenpakket. Diemen; 2002.
33. Stolk EA, Pickee SJ, Ament AHJA, Busschbach JJV. Equity in health care prioritisation: An empirical inquiry into social value. Health Policy. 2005;74(3):343–55.
34. Zorginstituut Nederland. Template Pharmacoeconomic dossier. Diemen; 2016.
35. World Health Organization. Report of the Commission on Macroeconomics and Health. Jakarta; 2002.
36. Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Econ. 2004;13(5):437–52.
37. Day BH. A meta-analysis of wage-risk estimates of the value of a statistical life. 1999.
38. Viscusi WK, Aldy JE. The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World. J Risk Uncertain. 2003;27(1):5–76.
39. Brouwer WBF. De basis van het pakket: Urgente uitdagingen voor de opzet en inzet van economische evaluaties in de zorg. Rotterdam: Instituut Beleid & Management Gezondheidszorg; 2009.

40. Zorginstituut Nederland. Uitvoeringstoets lage-ziektelastbenadering. Diemen; 2012.

41. Sen A. Inequality Reexamined. 2006th ed. Oxford University Press, editor. American Journal of Sociology. New York; 1992.

42. Rappange DR, Brouwer WBF. The evaluation of lifestyle interventions in the Netherlands. Health Econ Policy Law. 2012;7(2):243–61.

43. Bognar G. Fair innings. Bioethics. 2015;29(4):251–61.

44. Bognar G. Priority setting and age. In: Prioritization in medicine: an international dialogue. Switserland: Springer International Publishing; 2016. 163–177 p.

45. 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):703–14.

46. Office of Health Economics. Clarifying meanings of absolute and proportional shortfall with examples [Internet]. 2018 [cited 2018 Feb 6]. p. 1–6. Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/OHE-Note-on-proportional-versus-absolute-shortfall.pdf

47. Altmann S. Against proportional shortfall as a priority-setting principle. J Med Ethics. 2018;44(5):305–9.

48. Brazier J, Tsuchiya A. Improving Cross-Sector Comparisons: Going Beyond the Health-Related QALY. Appl Health Econ Health Policy. 2015;13(6):557–65.

49. Bobinac A, van Exel NJ., Rutten FFH, Brouwer WBF. The value of a QALY: Individual willingness to pay for health gains under risk. Pharmacoeconomics. 2014;32(1):75–86.

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

51. Brazier J, Rowen D, Mukuria C. Eliciting societal preferences for burden of illness, therapeutic improvement and end of life for value based pricing: A report of the main survey [Internet]. 2013 [cited 2017 Nov 10]. Available from: http://eprints.whiterose.ac.uk/99493/1/Eliciting societal preferences - 011.pdf

52. van de Wetering L, van Exel J, Bobinac A, Brouwer WBF. Valuing QALYs in Relation to Equity Considerations Using a Discrete Choice Experiment. Pharmacoeconomics. 2015;33(12):1289–300.

53. Bobinac A, van de Wetering L, van Exel NJA, Brouwer WBF. Equity-dependent social willingness to pay for a QALY. In: Valuing and refining outcome measures for economic evaluations in health care. 2015. p. 83–108.

54. Rowen D, Brazier J, Mukuria C, Keetharuth A, Hole AR, Tsuchiya A, et al. Eliciting societal preferences for weighting QALYs for burden of illness and end of life. Med Decis Mak. 2016;36(2):210–22.

55. 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.

56. Coast J, Smith R, Lorgelly P. Should the capability approach be applied in health economics? Health Econ. 2008;17:667–70.

57. 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):281–8.

58. National Institute for Health and Care Excellence. Consultation paper: Value based assessment of health technologies [Internet]. 2018. [cited 2017 Feb 22]. Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/VBA-TA-Methods-Guide-for-Consultation.pdf

59. National Institute for Health and Care Excellence. Responses to TA Methods Addendum Public Consultation between 27th March 2014 and 1st July 2014 [Internet]. 2014 [cited 2017 Feb 22]. Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/VBA-Consultation-Comments.pdf

60. National Institute for Health and Care Excellence. Value based assessment. Minutes of the Public Board Meeting of 17 September 2014, item 14/086. London; 2014.

61. Shah KK. Severity of illness and priority setting in healthcare: A review of the literature. Health Policy. 2009;93(2–3):77–84.

62. Lancsar E, Savage E. Deriving welfare measures from discrete choice experiments: Inconsistency between current methods and random utility and welfare theory. Health Econ. 2004;13(9):901–7.

2

Chapter 3
Equity weights for priority setting in healthcare: Severity, age, or both?

Based on: Reckers-Droog VT, van Exel NJA, Brouwer WBF. Equity weights for priority setting in healthcare: Severity, age, or both? Value in Health. 2019;22(12):1441-1449.

Abstract

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 operationalisations do not explicitly consider patients' age even though age may be inextricably related to severity and an equity-relevant characteristic. This study examines the relative strength of societal preferences for severity and age for informing allocation decisions in healthcare. We elicited preferences for severity and age in a representative sample of the public in the Netherlands (n=1,025) by applying choice- and person-trade-off (PTO) tasks in a design in which severity levels and ages varied both separately and simultaneously between patient groups. We calculated PTO ratios and, additionally, applied OLS regression models to aid interpretation of the ratios when severity and age both varied. 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. Our results suggest that severity operationalisations 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.

3.1 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 operationalised [3,4]. 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 [10,11]. Text box 3.1 includes a brief overview of the equity considerations and decision-making frameworks applied in these two countries.

Text box 3.1 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 operationalised as the disease-related loss of remaining QALYs without the new technology, compared to the QALY expectation in absence of the disease [10,12,13]. The Magnussen Committee suggested to divide AS into six severity classes and evaluate the ICERs of health technologies that target diseases with an AS of 0-3.9, 4-7.9, 8-11.9, 12-15.9, 16-19.9 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 (approximately €28,800 - €87,000), respectively [14,15]. Although the Norwegian government did not formally adopt the suggested AS classes and thresholds, they are informally used to inform allocation decisions in healthcare [10]. In the Netherlands, equity considerations are also explicated in terms of severity, and severity is defined in terms of proportional shortfall (PS). PS is operationalised 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) [11,16,17]. PS is divided into four severity classes. Health technologies that target diseases with a PS level of <0.10 are, in principle, not reimbursed [11,16]. The ICERs of health technologies that target diseases with PS levels of 0.10-0.40, 0.41-0.70, and 0.71-1.00 are evaluated against reference values of €20,000/QALY, €50,000/QALY, and €80,000/QALY, respectively [11,16,18–20]. 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 [11]. PH is a severity operationalisation that considers patients expected health and death and prioritises 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].

The decision-making frameworks applied in Norway and the Netherlands both account for societal preferences relating to the disease severity of patients. However, 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 prioritisation that favours 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 quality-adjusted life-years (QALYs), and hence may indirectly be prioritised 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, 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 [11,18]. However, older patients are, ceteris paribus, more likely to lose a larger fraction of their remaining QALYs, and hence may indirectly be prioritised over younger patients [13].

Although equity weights based on 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, prioritising 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 prioritising younger over older patients [26–34]. 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 [17,35]. 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 compare preferences for severity and age in relation to AS, PS, and prospective health (PH), as these severity operationalisations 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 operationalise these considerations in the context of healthcare priority setting.

3.2 Methods

3.2.1 Sample and data collection

A questionnaire was designed and, subsequently, distributed online in October 2018. Respondents (n=1,025) were quota sampled to represent the public in the Netherlands in terms of age (18-70 years), sex, and education level (see Table 3.1). Prior to collecting the data, we pilot tested the comprehensiveness of the questionnaire and clarity of the applied concepts and choice- and PTO tasks in two consecutive samples (n=120 and 1,023).

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 one of two patient groups (labelled 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 one year to standardise the health gains from treatment in different age groups. This avoided the influence of other considerations, including preferences for lifetime health and health maximisation. 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) and compared preferences for severity and age in relation to these severity operationalisations.

Table 3.1 Sample characteristics (n=1,014)[a]

	%	Mean (SD)
Age (Years)		45.3 (14.7)
Sex (Female)	51.7	
Education level[b]		
Low	8.3	
Medium	42.3	
High	49.4	
Health status (0-100 VAS)		75.4 (17.7)
Completion time of PTO tasks (in min)		7.7 (5.1)

PTO, person trade-off; VAS, visual analogue scale; [a] Respondents who completed the PTO tasks in less than 115 seconds (n=11) are excluded from this table; [b] Low = lower vocational and primary school; Medium = middle vocational and secondary school; High = higher vocational and academic education.

3.2.2 Questionnaire

The questionnaire consisted of three parts. In part one, we introduced three central concepts in text and graphs to respondents: (i) QOL, operationalised 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' [36]; (ii) severity, operationalised in terms of disease-related QOL loss; and (iii) treatment-related QOL gain. To familiarise respondents with these concepts and the applied tasks, we asked them to rate their current health on the VAS and complete one example choice- and one PTO task. We further asked them to assess the clarity of the concepts and example tasks on a seven-point Likert scale ranging from 1 'very unclear' to 7 'very clear'.

In part two, respondents stated their preferences for severity and age by completing a total of six randomly assigned choice- and associated PTO tasks in three 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 four iterative steps of how many patients (between 100 and 1,000,000) the other group should consist in order to be indifferent between the two groups. Figure 3.1 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 Appendix 3.1.

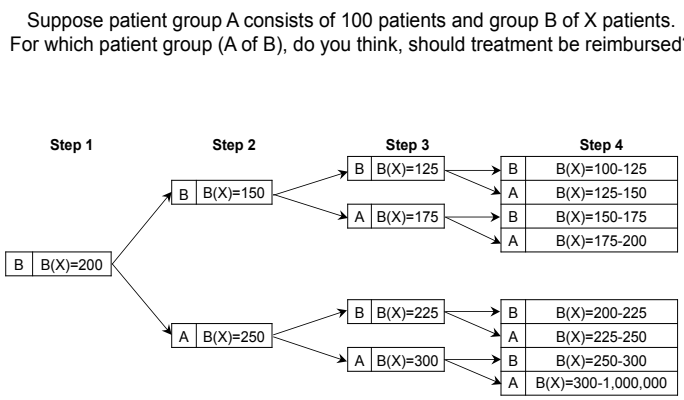


Fig. 3.1 Graphical representation of the iterative PTO process[a]

[a] In step four, 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.

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. Due to the disease, their QOL decreased from 100 to Y (with Y=20, 50, or 80) on the VAS for the duration of one 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 two choice- and associated PTO tasks based on a 'small' and 'large' difference in severity between the patient groups (i.e. 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 two choice- and associated PTO tasks based on a 'small' and 'large' difference in age between the groups (i.e. 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 two choice- and associated PTO tasks based on a 'small' and 'large' difference in both severity and age between the groups (i.e. 30 points/years and 60 points/years, respectively).

Table 3.2 presents the applied attributes and levels, and Table 3.3 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 three of the questionnaire, we asked respondents about their socio-demographic characteristics.

Table 3.2 Overview of attributes and levels

Attributes	Levels
Severity[a] (in points on VAS)	20; 50; 80
Age (in years)	10; 40; 70
QOL after treatment[b] (in points on VAS)	40; 70; 100
Treatment-related QOL gain (in points on VAS)	20
Life expectancy (in years)	80

QOL, health-related quality of life, VAS, visual analogue scale (range 0-100); [a] Severity is operationalised in terms of disease-related QOL loss; [b] QOL after treatment is calculated as: 100 – disease-related QOL loss + treatment-related QOL gain.

3.2.3 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 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 (i.e. per choice set) and pooled (i.e. per 'small' and 'large' difference) mean- and median-based PTO ratios, including our calculation methods, in Table 3.4. 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% CI) between ΔSeverity and ΔAge by bootstrapping (5,000 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 3.5. We focus the discussion of the results on the latter to account for outlying preferences. Note that, for calculating the preferences (in %) 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 '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 and a timed test of the tasks by three 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 (i.e. 1-3 level) for the clarity of the example PTO task and 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 [17,35], 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, Texas).

3.3 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 PTO task at 5.0 (1.7) on the seven-point Likert scale. Of the respondents, 177 (17.3%) reported a low score for the clarity of the example PTO task.

Table 3.3 presents the percentages of respondents with a preference for reimbursing treatment for group A, group B, or neither of the groups. Table 3.4 presents the PTO ratios.

In module 1, most respondents (range 47.4-58.9% across choice sets; Table 3.3) 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 3.4), the MOR was 2.51, indicating that respondents, on average, valued reimbursing treatment for one more severely ill patient equally to that for 2.51 less severely ill patients. The ROMs were relatively higher when the two patient groups were lower on the severity scale. The ROMs appeared to increase with age when patient groups were lower on the severity scale (i.e. choice sets 2, 4 and 6; Table 3.4). When the difference in severity was 'large' (choice sets 7 to 9; Table 3.4), 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 3.3) 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

Table 3.3 Preferences for reimbursing treatment for patient group A, B, or neither of the groups (in % of total) (n=1,014)[a]

	CS	Severity[b]		Age		Absolute shortfall[c]		Proportional shortfall[d]		Prospective health[e]			Preference (in %)		
		A	B	A	B	A	B	A	B	A	B	n	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, patient group A; B, patient group B; CS, choice set; QOL, health-related quality-of-life. [a] Respondents who completed the PTO tasks in less than 115 seconds (n=11) are excluded from this table; [b] Severity is operationalised in terms of disease-related QOL loss and measured in points (from 100) on a 0-100 visual analogue scale); [c] Absolute shortfall (AS) is calculated as ((disease-related QOL loss/100)*one year); [d] Proportional shortfall (PS) is calculated as ((disease-related QOL loss/100)*one year)/(80-age at onset of disease); [e] Prospective health is calculated as 80-age at onset of disease-absolute shortfall; * Choice set in which preferences are aligned with AS and PS.

1 to 6; Table 3.4), the MOR was 2.46, indicating that respondents, on average, valued reimbursing treatment for one younger patient equally to that for 2.46 older patients. The ROMs were relatively higher when it concerned comparisons between the higher age groups (i.e. choice sets 2, 4 and 6; Table 3.4). The ROMs did not show a clear relation with severity levels. When the difference was 'large' (choice sets 7 to 9; Table 3.4), 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 3.3) 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 3.4), the MOR was 1.98 and when this was 'large' (choice sets 9 to 10; Table 3.4), the MOR was 2.42.

Table 3.5 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 consisted 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 operationalisations is fully consistent with societal preferences when severity levels are equal and ages differ between patient groups. When severity levels and ages both vary, our results

Table 3.4 PTO results (n=1,014)[a]

		CS	Severity[b]		Age			Mean raw responses		Ratio of means	Median raw responses		Ratio of medians (ROMs)
			A	B	A	B	*n*	M_A	M_B	M_B / M_A	Mdn_A	Mdn_B	Mdn_B / Mdn_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	10190.43	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	10622.85	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						1,014						
	Ratio of means							599.24	7558.40	12.61			
	Median of ratios (MORs)												2.51
		7	80	20	10	10	336	545.00	13803.70	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	12814.76	18.64	100.00	275.00	2.75
	n						1,014						
	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	15797.84	85.69	100.00	301.00	3.01
		5	20	20	10	40	171	224.25	11189.68	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						1,014						
	Ratio of means							313.23	7567.21	24.16			
	Median of ratios (MORs)												2.46

		CS	Severity[b]		Age			Mean raw responses		Ratio of means	Median raw responses		Ratio of medians (ROMs)
			A	B	A	B	*n*	M_A	M_B	M_B / M_A	Mdn_A	Mdn_B	Mdn_B / Mdn_A
		7	80	80	10	70	338	128.17	12564.78	98.03	100.00	400.00	4.00
		8	50	50	10	70	341	474.91	13564.93	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						1,014						
	Ratio of means							244.28	10369.07	42.45			
	Median of ratios (MORs)												3.50
Module 3		1	80	50	10	40	129	224.66	18543.20	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	18591.40	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	18575.50	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						1,014						
	Ratio of means							1174.65	9067.36	7.72			
	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						1,014						
	Ratio of means							2715.76	9447.29	3.48			
	Median of ratios (MORs)												2.42

A, patient group A; B, patient group B; CS, choice set; QOL, health-related quality of life; [a] Respondents who completed the PTO tasks in less than 115 seconds (n=11) are excluded from this table; [b] Severity is operationalised in terms of disease-related QOL loss (in points on a 0-100 visual analogue scale).

Table 3.5 Regression results[a]

	Model 1[b]			
DV:	Including outliers		Excluding outliers	
∆nrPatients*QOLgain	β (SE)	95% CI	β (SE)	95% CI
∆Severity	5245.40** (1662.67)	1982.73, 8508.07	1193.73*** (324.82)	556.33, 1831.14
∆Age	-2772.48· (1662.67)	-6035.15, 490.19	-968.02** (324.82)	-1605.42, -330.61
R^2	0.01		0.02	
n	1,014		1,005	

	Model 2[c]			
DV:	Including outliers		Excluding outliers	
∆nrPatients*QOLgain	β (SE)	95% CI	β (SE)	95% CI
∆Severity	2215.53** (791.01)	663.32, 3767.73	985.00*** (272.28)	450.69, 1519.31
∆Age	-2051.20** (791.01)	-3603.41, -498.99	-1490.30*** (272.28)	-2024.61, -956.00
R^2	0.01		0.04	
n	1,014		1,006	

DV, dependent variable; SE, standard error; QOL gain, treatment-related gain in quality of life (20 points on visual analogue scale); ∆Age, difference in age (in years) between patient group A and B; ∆nrPatients*QOLgain, difference in the number of patients between patient group B and A, each multiplied by the treatment-related gain in QOL; ∆Severity, difference in severity (in points QOL loss) between patient group A and B; DV, dependent variable; [a] Respondents who completed the PTO tasks in less than 115 seconds (n=11) are excluded from this analysis. The results are presented in- and excluding respondents with outlying (-1.96> z-score >1.96) preferences; [b] Model 1 presents the results for a difference of 30 points in QOL loss and 30 years in age between the patient groups; [c] Model 2 presents the results for a difference of 60 points in QOL loss and 60 years in age between the patient groups; · p<0.10; * p<0.05; ** p<0.01; *** p<0.001.

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).

3.4 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 and/or older patient group. This preference heterogeneity is consistent with the results of previous studies [4].

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 is fully consistent with societal preferences when severity levels are equal and ages differ between patient groups. When severity levels and ages both 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. [35], 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 [24,26–35]. For example, our results are consistent with those of Dolan and Tsuchiya [27] who found that preferences for age were stronger than those for severity when preferences for both were elicited simultaneously. 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, e.g. discrete choice experiments, 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 two 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 one year. We are aware that this comes with the limitation that our results cannot be generalised 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 [4,27,40]. 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 com-

bination 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 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 like spill-over effects on family or informal carers, and productivity losses [29,32]. 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 upwards 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 downwards by this. A final limitation concerns the low R-squared values of the regression models. These low values indicate that the models do not explain much of the data variability. However, the implied ratios associated with an increase of 30 points and 30 years increase in the difference in severity and age between the patient groups resemble the mean-based PTO ratios in Table 3.4 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 [20,26–30,35]. However, our results also suggest that none of these severity operationalisations 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 operationalisations that do not indirectly favour younger over older patients, like AS does. Whether or not age weights are 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 paper.

Further research is warranted to examine the relative strength of preferences for severity and age when severity is operationalised on a QALY scale, thus incorporating both length and quality of life 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 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 weights for informing allocation decisions in healthcare.

3.5 Conclusions

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

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 Health Econ. 2013;14(1):107–15.
2. Ottersen T. Lifetime QALY prioritarianism in priority setting. J Med Ethics. 2013;39(3):175–80.
3. Hernæs UJV, Johansson KA, Ottersen T, Norheim OF. Distribution-Weighted Cost-Effectiveness Analysis Using Lifetime Health Loss. Pharmacoeconomics. 2017;35(9):965–74.
4. Shah KK. Severity of illness and priority setting in healthcare: A review of the literature. Health Policy. 2009;93(2–3):77–84.
5. Schwappach DL. Resource allocation, social values and the QALY: A review of the debate and empirical evidence. Health Expect. 2002;5(3):210–22.
6. Sabik LM, Lie RK. Priority setting in health care: Lessons from the experiences of eight countries. Int J Equity Health. 2008;7:1–13.
7. Nord E. The person-trade-off approach to valuing health care programs. Med Decis Mak. 1995;15(3):201–8.
8. Williams A. Intergenerational equity: An exploration of the "fair innings" argument. Health Econ. 1997;6(2):117–32.
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. Norwegian Ministry of Health and Care Services. Principles for priority setting in health care. Oslo; 2017.
11. Zorginstituut Nederland. Ziektelast in de praktijk. Diemen; 2018.
12. Ottersen T, Førde R, Kakad M, Kjellevold A, Melberg HO, Moen A, et al. A new proposal for priority setting in Norway: Open and fair. Health Policy. 2016;120(3):246–51.
13. Statens Legemiddelverk. Beregning av absolutt prognosetap i hurtig metodevurderinger av legemidler – bruke beregningsprinsippet for behandling eller forebygging ? [Internet]. 2019. Available from: https://legemiddelverket.no/Documents/Offentlig finansiering og pris/Dokumentasjon til metodevurdering/APT- behandling eller forebygging_310119.pdf
14. Magnussen J, Aaserud M, Granaas T, Al. E. På ramme alvor—Alvorlighet og prioritering. [Internet]. 2015. Available from: https://www.regjeringen.no/contentassets/d5da48ca5d1a4b128c72fc5daa3b4fd8/paa_ramme_alvor.pdf
15. Magnussen J. Severity of illness and priority setting in Norway: summary of a report from a working group [Internet]. 2015 [cited 2020 Jan 13]. Available from: https://www.regjeringen.no/contentassets/d5da48ca5d1a4b128c72fc5daa3b4fd8/summary_the_magnussen_report_on_severity.pdf
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):621–9.

17. 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):1097–107.
18. Stolk EA, Goes E, Kok E, Van Busschbach JJV. Uitwerking criteria noodzakelijkheid, eigen rekening en verantwoording en lifestyle. In: Breedte geneesmiddelenpakket. Amstelveen: College voor Zorgverzekeringen; 2001.
19. Versteegh MM, Ramos IC, Buyukkaramikli NC, Ansaripour A, Reckers-Droog VT, Brouwer WBF. Severity-Adjusted Probability of Being Cost Effective. Pharmacoeconomics. 2019;37(9):1155–63.
20. Zorginstituut Nederland. Cost-effectiveness in practice. Diemen; 2015.
21. Dolan P, Olsen JA. Equity in health: The importance of different health streams. J Health Econ. 2001;20(5):823–34.
22. Brazier J, Rowen D, Mukuria C. Eliciting societal preferences for burden of illness, therapeutic improvement and end of life for value based pricing: A report of the main survey [Internet]. 2013 [cited 2017 Nov 10]. Available from: http://eprints.whiterose.ac.uk/99493/1/Eliciting societal preferences - 011.pdf
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(2–3):281–8.
24. Nord E. Concerns for the worse off: Fair innings versus severity. Soc Sci Med. 2005;60(2):257–63.
25. Edlin R, Round J, McCab C, Sculpher M, Claxton K, Cookson R. Cost-effectiveness analysis and ageism: A review of the theoretical literature. A report for the department of health. Leeds: University of Leeds; 2008.
26. 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.
27. 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):703–14.
28. Olsen JA. Priority preferences: "End of life" does not matter, but total life does. Value Health. 2013;16(6):1063–6.
29. Bobinac A, van de Wetering L, van Exel NJA, Brouwer WBF. Equity-dependent social willingness to pay for a QALY. In: Valuing and refining outcome measures for economic evaluations in health care. 2015. p. 83–108.
30. 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.
31. 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):1–18.
32. 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):1119–26.

33. Johannesson M, Johansson PO, Johanesson M, Johanesson P. Is the valuation of a QALY gained independent of age? Some empirical evidence. J Health Econ. 1997;16(5):589–99.
34. 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):687–96.
35. Stolk EA, Pickee SJ, Ament AHJA, Busschbach JJV. Equity in health care prioritisation: An empirical inquiry into social value. Health Policy. 2005;74(3):343–55.
36. Oemar M, Janssen B. EQ-5D-5 L user guide-basic information on how to use the EQ-5D-5 L instrument. Rotterdam: EuroQol Group; 2013.
37. Johansson-Stenman O, Martinsson P. Are some lives more valuable? An ethical preferences approach. J Health Econ. 2008;27(3):739–52.
38. Lancsar E, Savage E. Deriving welfare measures from discrete choice experiments: Inconsistency between current methods and random utility and welfare theory. Health Econ. 2004;13(9):901–7.
39. van de Wetering L, van Exel J, Bobinac A, Brouwer WBF. Valuing QALYs in Relation to Equity Considerations Using a Discrete Choice Experiment. Pharmacoeconomics. 2015;33(12):1289–300.
40. Ubel PA. How stable are people's preferences for giving priority to severely ill patients? Soc Sci Med. 1999;49(7):895–903.
41. van de Wetering EJ, van Exel NJA, Rose JM, Hoefman RJ, Brouwer WBF. Are some QALYs more equal than others? Eur J Health Econ. 2016;17(2):117–27.
42. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science. 1981;211(4481):453–8.
43. Dolan P. The measurement of individual utility and social welfare. J Health Econ. 1998;17(1):39–52.
44. Brouwer WBF, Van Exel NJA, Stolk EA. Acceptability of less than perfect health states. Soc Sci Med. 2005;60(2):237–46.
45. Doctor JN, Miyamoto JM. Person tradeoffs and the problem of risk. Expert Rev Pharmacoeconomics Outcomes Res. 2005;5(6):677–82.
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):1–10.
47. Ternent L, Tsuchiya A. A note on the expected biases in conventional iterative health state valuation protocols. Med Decis Mak. 2013;33(4):544–6.

Appendix 3.1: Person trade-off approach and task examples

Person trade-off approach

Respondents completed a total of six choice- and six associated person trade-off (PTO) tasks across three similarly structured modules, thus completing two choice- and two PTO tasks in each module. Before completing the first choice task in a module, we assigned respondents to a choice set in which the attributes and levels represented a 'small' difference of 30 points in QOL loss and/or 30 years in age between the patient groups. Based on this difference and given the contingency that both groups consisted of 100 patients, we asked respondents to state a preference for reimbursing treatment for patient group A or B, or, in case of indifference for neither of the groups. We then presented the same choice set to respondents and explained that the groups now differed in size. We set respondents' patient group of preference in the preceding choice task (or, in case of no preference, a randomly selected group) as reference group and explained that this group again consisted of 100 patients. We then asked respondents to complete a PTO task in which they could state in three iterative steps (see Figure 3.1 for a graphical representation of the iterative PTO process, including the steps taken and the interval between the steps) of how many patients (between 100 and 1,000,000) the other group should consist in order for them to be indifferent. Subsequently, we assigned respondents to a second choice set in which the attributes and levels represented a 'large' difference of 60 points in QOL loss and/or 60 years between the patient groups, based on which we asked them to complete a second choice- and PTO task, similar in set up as the first.

Choice task example

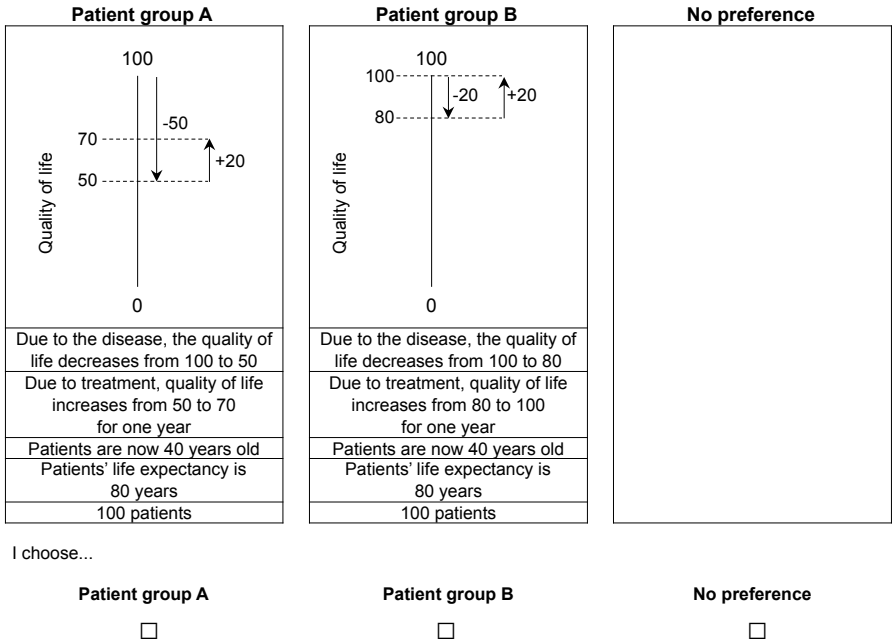
The disease-related loss in quality of life can differ between patients. In the graph below, you see two patient groups that would have had a life expectancy of 80 years in full health had they not become ill. Both patient groups have become ill at the age of 40. The disease lasts for one year and leads to a lower quality of life during this year.

Due to the disease, the quality of life of patients in group A decreases from 100 to 50 and in group B from 100 to 80 on a scale from 0 (the worst health you can imagine) to 100 (the best health you can imagine). Treatment can reduce the effects of the disease. As a result, the quality of life in both groups during the year is 20 points higher than without treatment. In group A, the quality of life will increase from 50 to 70 and in group B from 80 to 100. After this year, all patients will regain their full health.

There are no other differences between the two patient groups. The treatments and costs are also the same.

Imagine that the treatment of only one of the two patient groups can be reimbursed from the public health insurance package and that you could advise health policy makers on what would be the optimal way of allocating the healthcare budget.

If both patient groups consist of 100 patients, for which patient group, do you think, should the treatment be reimbursed?



In the example of the PTO task below, we assume that the respondent has stated a preference for patient group A in the preceding choice task, and hence the subsequent PTO task is set at 100 patients for patient group A.

PTO task example

The disease-related loss in quality of life can differ between patients. In the graph below, you see the same two patient groups that would have had a life expectancy of 80 years in full health had they not become ill. Both patient groups have become ill at the age of 40. The disease lasts for one year and leads to a lower quality of life during this year.

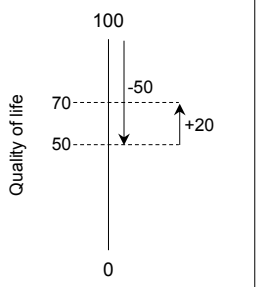
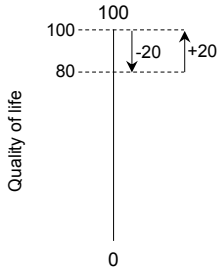
Due to the disease, the quality of life of patients in group A decreases from 100 to 50 and in group B from 100 to 80 on a scale from 0 (the worst health you can imagine) to 100 (the best health you can imagine). Treatment can reduce the effects of the disease. As a result, the quality of life in both groups during the year is 20 points higher than without treatment. In group A, the

quality of life will increase from 50 to 70 and in group B from 80 to 100. After this year, all patients will regain their full health.

There are no other differences between the two patient groups. The treatments and costs are also the same.

Imagine that the treatment of only one of the two patient groups can be reimbursed from the public health insurance package and that you could advise health policy makers on what would be the optimal way of allocating the healthcare budget.

You have stated a preference for reimbursing treatment for patient group A if both groups consisted of 100 patients. Now, suppose that the two patient groups differ with regard to the number of patients in each group.

Patient group A	Patient group B
	
Due to the disease, the quality of life decreases from 100 to 50	Due to the disease, the quality of life decreases from 100 to 80
Due to treatment, quality of life increases from 50 to 70 for one year	Due to treatment, quality of life increases from 80 to 100 for one year
Patients are now 40 years old	Patients are now 40 years old
Patients' life expectancy is 80 years	Patients' life expectancy is 80 years
100 patients	? patients

In the four steps below, we assume that the respondent stated a preference for reimbursing treatment for patient group A in the choice task, for patient group A in step one, for patient group B in step two, and again for patient group B in step three of the PTO process.

Step one

Suppose that group A consists of 100 patients and group B of 200 patients. For which patient group, do you think, the treatment should be reimbursed?

■ Patient group A

☐ Patient group B

Step two

Suppose that group A consists of 100 patients and group B of 250 patients. For which patient group, do you think, the treatment should be reimbursed?

☐ Patient group A

■ Patient group B

Step three

Suppose that group A consists of 100 patients and group B of 225 patients. For which patient group, do you think, the treatment should be reimbursed?

☐ Patient group A

■ Patient group B

Step four

You have stated a preference for reimbursing the treatment for patient group A if group A consists of 100 patients and group B of 200 patients, but your preference shifts to patient group B if group A consists of 100 patients and group B of 225 patients.

Of how many patients should group B consist (more than 200 and less than 225) in order for your preference for patient group A and B to be equally strong?

Note: For respondents with a consistent preference for one of the patient groups, the range in step four was restricted to a maximum of 1,000,000 patients.

3

Chapter 4
Willingness to pay for health-related quality of life gains in relation to disease severity and the age of patients

Based on: Reckers-Droog VT, van Exel NJA, Brouwer WBF. Willingness to pay for health-related quality of life gains in relation to disease severity and the age of patients. Forthcoming in Value in Health. 2021.

Abstract

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. We obtained WTP estimates by applying contingent-valuation tasks in a representative sample of the public in the Netherlands (n=2,023). We applied random-effects Generalised Least Squares regression models to estimate the effect of patients' disease severity and age, size of QOL gains, disease outcome (full recovery/death one year after falling ill), and respondent characteristics on the WTP. 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 non-linear 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. 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.

4.1 Introduction

An important objective of publicly financed healthcare systems is to maximise 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 (e.g. 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 4.1 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 4.1) are increasingly studied, also in relation to patients' age. The available evidence 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 [1,5,11,40–43]. Nevertheless, the weights applied in Norway, the Netherlands, and—classifying terminally ill patients as severely ill [44]—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 prioritise younger over older patients and, conversely, the weights based on proportional shortfall in

Text box 4.1 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) gained 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 (i.e. an absolute shortfall of ≥20 QALYs) [31,37,38]. 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 [31,39,40]. 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 (i.e. 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 (i.e. a proportional shortfall of <0.10) are generally not recommended for reimbursement [28,41,42]. 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 three months [43]. 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) [43].

the Netherlands and end-of-life considerations in England may implicitly prioritise older over younger patients [11,45–47].

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 (one year after falling ill). Based on the available evidence on societal preferences in this context, we hypothesised that WTP would be higher for more severely ill and younger patients, and for patients who can be considered terminally ill. We further hypothesised that the elicited WTP would be sensitive to scale and to household income, indicating the theoretical validity of the elicited WTP [48].

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 the Netherlands, Norway, 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.

4.2 Methods

4.2.1 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=2,023).

Before respondents completed the questionnaire, we explained that healthcare resources are scarce and decision makers use information on societal preferences in order 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 [49]. 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 [49], 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 [49,50]. Upon completion of the questionnaire, respondents received a fee of 50 eurocents that they could save in a personal account or donate to charity.

4.2.2 Questionnaire

The questionnaire consisted of four parts. In part one, we introduced respondents to the following concepts using text and graphs: (i) QOL, operationalised on a visual analogue scale (VAS) ranging from 0 "dead" to 100 "full health", (ii) disease severity, operationalised 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 one year after falling ill, and (iii) treatment-related QOL gain (in points on the VAS). We familiarised respondents with the concepts and tasks by asking them to assess their own QOL 'today' on the VAS and complete a practice task from an individual perspective [49]. After completing this task, we asked respondents to assess its level of clarity on a seven-point Likert scale ranging from 1 "very unclear" to 7

"very clear" and to indicate on what they would likely economise to cover the stated WTP to increase their awareness of the associated opportunity costs.

In part two, respondents completed two tasks from a SIP perspective for which they were randomly assigned to two 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 one year. After this year, they would fully recover (i.e. 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 one 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 two-step procedure described in Text box 4.2. After respon-

Text box 4.2 Contingent-valuation procedure

Willingness to pay was elicited by applying a two-step contingent-valuation procedure, consisting of a payment scale and a bounded open-ended question [51]. The payment scale ranged from a €0 to €24 increase in monthly basic health-insurance premium with unevenly distributed intervals between the value points (i.e. €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 all adults (18+) in the Netherlands. By approximation, the number of adults was 13.7 million and the monthly premium €115.00 per person in 2019 [52]. In step one, 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 one 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 one year. In step two, we asked respondents to indicate the maximum increase in monthly premium they would be willing to pay within the range obtained in step one. 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 one to explain their main reason for having this preference by completing an open-text field or checking one of six randomised answer options. Three answer options related to true zero valuations (i.e. "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 three answer options related to protest zero valuations (i.e. "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 two authors and subsequently classified as a true or protest zero valuation.

dents 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 seven-point Likert scale ranging from 1 "very uncertain" to 7 "very certain".

In part three, respondents completed one additional task from a SIP perspective for which they were randomly assigned to one out of 20 scenarios that were evenly distributed across five 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) one 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 [53]. 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 4.1 includes a task example and Appendix 4.2 an overview of the scenario characteristics.

In part four, we asked respondents about their socio-demographic characteristics.

4.2.3 Statistical analyses and hypotheses

Before conducting the analyses, we excluded protest zero valuations (see Text box 4.2), outliers, and speeders. We identified outliers based on the distribution of stated WTPs (z-score ≥1.64). We classified respondents who completed the three tasks in less than 90 seconds as speeders, based on a timed test of completing the tasks by three 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 died one year after falling ill. We applied two-tailed Welch's t-tests (Bonferroni corrected) to examine whether the latter was statistically significantly different from 0. Furthermore, we applied seven random-effects Generalised 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, one 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, age2, 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 [54]), 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 (labelled 'CV task') [55]. A downwards 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 [56]. 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 (i.e. 1–3 level) for the practice task, reported a low certainty score (i.e. 1–3 level) for actually paying the stated WTP, and completed the three 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 hypothesised that respondents' WTP would be higher for QOL gains in more severely ill patients (i.e. patients with a higher level of disease severity and patients who died one year after falling ill) and for QOL gains in younger patients. Moreover, we hypothesised 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 [48].

We conducted the analyses using Stata 16.1 (Stata Corp LP, College station, Texas).

4.3 Results

Table 4.1 presents the descriptive statistics of the sample (n=1,317) 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 seven-point Likert scales. A total of 50 (3.8%) respondents reported a low

Table 4.1 Descriptive statistics

	Sample (n=1,317)[a]		General public[b]	
	%	Mean (SD)	%	Mean
Age (Years)		51.5 (15.9)		46.1
Sex (Female)	51.3		50.3	
Education level[c]				
Low	11.6		8.6	
Medium	58.3		57.4	
High	30.1		32.5	
Household income (After tax)				
<€1,999	31.7			
€2,000 – €3,999	42.1			
≥€4,000	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, contingent valuation; NS, Not Stated; QOL, health-related quality of life; VAS, visual analogue scale (ranging from 0 "dead" to 100 "full health"); a 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; b Age is based on statistics for population aged 18-75 years, sex is based on statistics for the overall population, and education level is based on statistics for population aged 15-75 years. Population statistics for 2019, source: Statistics Netherlands (https://opendata.cbs.nl/statline); c Low = lower vocational and primary school, Medium = middle vocational and secondary school, High = higher vocational and academic education.

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 4.2 presents the raw mean (SD; 95% CI) WTP for QOL gains of 10 points in patients who fully recovered one year after falling ill. On average, the WTP was €8.0 per month for the duration of one 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 non-linear. 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. As 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 4.3 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 one 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 4.2). These results indicate that respondents' WTP was generally higher for larger sized QOL gains and 'hump shaped' across ages,

Table 4.2 Raw mean (SD; 95% CI) WTP for QOL gains of 10 points in patients who will fully recover one year after falling ill (in €)[a]

	Severity[b]										
	10		30		50		70		90		Average
Age[c]	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)	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, health-related quality of life; WTP, willingness to pay (in € per month for the duration of one year); [a] 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 one year after falling ill. The treatment-related QOL gain is 10 points, measured on a VAS ranging from 0 “death” to 100 “full health”; [b] Severity is operationalised in terms of disease-related QOL loss and measured in points from 100 on the VAS; [c] Age at onset of the disease (in years).

Table 4.3 Raw mean (SD; 95% CI) WTP for QOL gains in patients who will fully recover and die one year after falling ill (in €)[a]

	Disease outcome one year after falling ill											
	Full recovery						Death				Difference	
				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
Age[b]	n	Mean	n	Mean	n	Mean	n	Mean	n	Mean	ΔMean	ΔMean
		(SD; 95% CI)		(SD; 95% CI)		(SD; 95% CI)		(SD; 95% CI)		(SD; 95% CI)	(SE; 95% CI)	(SE; 95% CI)
10	134	8.2	66	8.5	59	10.4	61	7.5	71	9.8	-1.0	-0.6
		(6.4; 7.1, 9.30		(6.1; 7.0, 10.00		(11.0; 7.6, 13.3)		(6.2; 5.9, 9.1)		(6.7; 8.2, 11.4)	(1.1; -3.2, 1.2)	(1.6; -3.9, 2.6)
20	121	7.3	61	8.9	66	12.3	57	9.7	64	7.8 (6.2; 6.3, 9.4)	0.7	-4.5*
		(6.7; 6.1, 8.50		(7.1; 7.1, 10.8)		(12.9; 9.2, 15.5)		(7.4; 7.7, 11.6)			(1.3; -1.9, 3.4)	(1.8; -8.0, -1.0)
40	129	8.5	61	7.4	61	10.1	59	7.9	61	8.9	0.6	-1.2
		(7.1; 7.2, 9.70		(5.8; 5.9, 8.9)		(10.1; 7.5, 12.6)		(7.5; 6.0, 9.9)		(8.1; 6.8, 10.9)	(1.2; -1.9, 3.0)	(1.7; -4.5, 2.1)
70	128	7.8	54	7.5	68	8.6	61	5.6	60	8.0	-1.9	-0.6
		(7.1; 6.5, 9.00		(6.4; 5.7, 9.2)		(9.4; 6.3, 10.8)		(5.1; 4.3, 6.8)		(6.9; 6.2, 9.8)	(1.1; -4.1, 0.2)	(1.4; -3.4, 2.3)

QOL, health-related quality of life (measured in points on a visual analogue scale (VAS), ranging from 0 "dead" to 100 "full health"); WTP, willingness to pay (in € per month for the duration of one year); [a] 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 VAS; [b] Age at onset of the disease (in years); * $p<0.05$ (Bonferroni corrected, $\alpha=0.05/4$).

Table 4.4 Regression results[a]

	Model 1		Model 2		Model 3		Model 4[b]	
DV: WTP	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
Severity[c] (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, 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[d] (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)[e]					-0.80* (0.34)	-1.47, -0.13	-0.79* (0.34)	-1.46, -0.12
CV task	-0.83*** (0.12)	-1.07, -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)	2,521; 1,317		2,251; 1,317		3,443; 1,308		3,443; 1,308	

CV, 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 one year); [a] 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 two tasks in less than 60 seconds (based on timed test of completing the tasks by three 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; [b] 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; [c] Severity is operationalised in terms of disease-related QOL loss and measured in points from 100 on the VAS; [d] Age at onset of the disease (in years); [e] Disease outcome one year after falling ill; $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Table 4.5 Regression results[a]

	Model 5[b]		Model 6[b]		Model 7[b]	
DV: WTP	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
Severity level[c] (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[d] (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)[e]	-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				

	Model 5[b]		Model 6[b]		Model 7[b]	
DV: WTP	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
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
Age^2 (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[f] (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[g] (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
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)[g]	3,250; 1,235		3,250; 1,235		3,250; 1,235	

CV, 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 one year); [a] Note that models 5-7 are based on the data obtained in part 2 and 3 of the questionnaire; [b] 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; [c] Severity is operationalised in terms of disease-related QOL loss and measured in points from 100 on the VAS; [d] Age at onset of the disease (in years); [e] Disease outcome one year after falling ill; [f] Low = lower vocational and primary school, Medium = middle vocational and secondary school, High = higher vocational and academic education; [g] Household income is adjusted for household size and calculated as household income / (household size)^0.5; [g] Note that that the number of observations and groups is lower than in Table 4.4 as respondents who did not state their household income are excluded from Table 5; · p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

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, α/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 4.4 and 4.5 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. 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 (e.g. 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 onwards (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 onwards (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 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 one 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 as this was dependent on patients' level of disease severity in some scenarios.

The results presented in Table 4.5 provide further insight into the interactions between patients' disease severity, the size of QOL gains, and disease outcome and age of patients (see Figure 4.1), and into the effect of respondent characteristics on WTP. The results indicate that respondents' WTP for

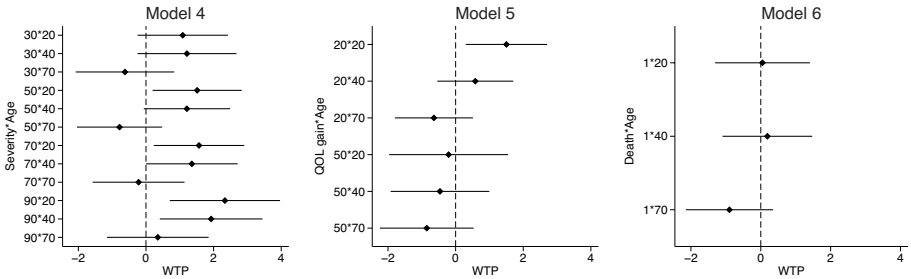


Fig. 4.1 Graphical presentation of interaction terms (mean additional effect; 95% CI) presented in Table 4.5

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 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.

4.4 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 one 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 one year after falling ill. However, the relations were non-linear 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 onwards. 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 [45,57–59].

We would like to make four remarks in relation to our findings. Firstly, we applied a SIP perspective for eliciting respondents' WTP and, 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 (e.g. associated with having children and respondents' proximity to the age of patients), we acknowledge that unobservable self-regarding preferences (e.g. associated with the probability of respondents' own need for treatment) may have impacted our results. Secondly, 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 societal perspective that excludes respondents from the hypothetical patient group [49]. Thirdly, we observed considerable preference heterogeneity, which is consistent with the findings of other studies that examined societal preferences in this context [1,10,12,53]. 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 one year. Hence, on an aggregated level the (differences in) WTP per QALY is substantial. The treatment generates 1000 QALYs (i.e. 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 4.1) and likely influenced by the scenario characteristics (e.g. the number of patients and certainty of QOL gains), it should be noted that such high values are not uncommon in preference-elicitation studies [61] and, considering the high ICERs of some health technologies that are currently reimbursed in the Netherlands [62,63], 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 [11,40,64–66] and to larger sized health gains [53], and with those of other studies that find no support for attaching a higher weight to gains in terminally ill patients [67–70]. 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' [e.g. 60–63]. Although there

may be a moral case for attaching a higher weight to health gains in terminally ill patients [75], our findings—like those of previous studies [67–70]—suggest that empirical support for applying a higher weight to these gains may be limited. This is recognised 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 4.1) by considerations that relate more broadly to the disease severity of patients in order to better align their decision-making framework with societal preferences [75,76].

However, as 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 societal 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 operationalised disease severity and the health gain in the specific scenarios, i.e. 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 two-step CV procedure. Although we could have applied other methods (e.g. 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 randomisation of scenarios, exclusion of speeders, restriction of the disease duration to one year, and standardisation of patients' risk of falling ill and dying within a certain time frame (i.e. 100%, implying no uncertainty), the size of the patient group and QOL gains as well as the costs of treatment as this reduced the possible influence of an order effect, satisficing behaviour [77], cognitive biases associated with risk assessment [78], and of other considerations (e.g. related to health maximisation 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 generalised 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 time frame (i.e. 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 [79]. 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 [80,81]. However, we cannot rule out the possibility that the scale influenced respondents' WTP, particularly in case of unstable or not (yet) well-formed preferences [80]. 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 [82]. 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 s [83]. 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 behaviour of respondents may have offered a counterbalance [84]. 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 [85]. A final limitation concerns the low R2 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 variance as possible. Rather, our aim was to assess whether WTP was influenced by scenario characteristics (i.e. 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 [69,86–89]. 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.

As there is much variation between the results and designs of studies that examine the strength of societal preferences [60], 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 societal or SIP perspective.

4.5 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 non-linear 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.

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 Health Econ. 2013;14(1):107–15.
2. Weinstein MC, Torrance G, McGuire A. QALYs: The basics. Value Health. 2009;12:S5–9.
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. In: Valuing and refining outcome measures for economic evaluations in health care. 2015. p. 83–108.
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):289–90.
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):1097–107.
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):157–71.
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):1119–26.
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):1441–9.
12. Round J, Paulden M. Incorporating equity in economic evaluations: a multi-attribute equity state approach. Eur J Health Econ. 2018;19(4):489–98.
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. Health Expect. 2002;5(3):210–22.
15. Donaldson C, Baker R, Mason H, Jones-Lee M, Lancsar E, Wildman J, 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):208–38.
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):1163–72.

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;126:128–37.
20. Norheim OF, Baltussen R, Johri M, Chisholm D, Nord E, Brock DW, 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. Daniels. Accountability 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;14:255–63.
24. Daniels N, Sabin JE. Accountability for reasonableness: an update. BMJ. 2008;337.
25. Chalkidou K, Glassman A, Marten R, Vega J, Teerawattananon Y, Tritasavit N, et al. Priority-setting for achieving universal health coverage. Bull World Health Organ. 2016;94(6):462–7.
26. Broqvist M, Sandman L, Garpenby P, Krevers B. The meaning of severity – do citizenś views correspond to a severity framework based on ethical principles for priority setting? Health Policy. 2018;122(6):630–7.
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 2013. London; 2013.
31. Norwegian Ministry of Health and Care Services. Principles for priority setting in health care. Oslo; 2017.
32. 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):703–14.
33. Stolk EA, Pickee SJ, Ament AHJA, Busschbach JJV. Equity in health care prioritisation: An empirical inquiry into social value. Health Policy. 2005;74(3):343–55.
34. Olsen JA. Priority preferences: "End of life" does not matter, but total life does. Value Health. 2013;16(6):1063–6.
35. 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.
36. Cookson R. Can the nice "end-of-life premium" be given a coherent ethical justification? J Health Polit Policy Law. 2013;38(6):1129–48.
37. Magnussen J, Aaserud M, Granaas T, Al. E. På ramme alvor—Alvorlighet og prioritering. [Internet]. 2015. Available from: https://www.regjeringen.no/contentassets/d5da48ca5d1a4b128c72fc5daa3b4fd8/paa_ramme_alvor.pdf

38. Magnussen J. Severity of illness and priority setting in Norway: summary of a report from a working group [Internet]. 2015 [cited 2020 Jan 13]. Available from: https://www.regjeringen.no/contentassets/d5da48ca5d1a4b128c72fc5daa3b4fd8/summary_the_magnussen_report_on_severity.pdf
39. Ottersen T, Førde R, Kakad M, Kjellevold A, Melberg HO, Moen A, et al. A new proposal for priority setting in Norway: Open and fair. Health Policy. 2016;120(3):246–51.
40. Statens Legemiddelverk. Beregning av absolutt prognosetap i hurtig metodevurderinger av legemidler – bruke beregningsprinsippet for behandling eller forebygging ? [Internet]. 2019. Available from: https://legemiddelverket.no/Documents/Offentlig finansiering og pris/Dokumentasjon til metodevurdering/APT- behandling eller forebygging_310119.pdf
41. Zorginstituut Nederland. Uitvoeringstoets lage-ziektelastbenadering. Diemen; 2012.
42. 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):621–9.
43. National Institute for Health and Care Excellence. Guide to the Methods of Technology Appraisal. London; 2013.
44. Williams A. Intergenerational equity: An exploration of the "fair innings" argument. Health Econ. 1997;6(2):117–32.
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 [Internet]. October 10. 2019 [cited 2020 Apr 1]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/mortality/age#heading-Zero
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):158–68.
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):545–51.
49. Bobinac A, Van Exel NJA, Rutten FFH, Brouwer WBF. Valuing qaly gains by applying a societal perspective. Health Econ. 2013;22(10):1272–81.
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):1046–55.
51. Ministry of Health, Welfare and Sport. Wat kost het om een zorgverzekering af te sluiten? [Internet]. [cited 2020 Jan 13]. Available from: https://www.rijksoverheid.nl/onderwerpen/zorgverzekering/vraag-en-antwoord/premie-zorgverzekering
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):453–8.

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):313–29.

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;181:158–67.

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 Health Econ. 2019;20:31–41.

58. 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):281–8.

59. 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).

60. Rombach S, Hollak C, Linthorst G, Dijkgraaf M. Cost-effectiveness of enzyme replacement therapy for Fabry disease. Orphanet J Rare Dis. 2013;8(1):1–9.

61. Kanters TA, van der Ploeg AT, Kruijshaar ME, Rizopoulos D, Redekop WK, Rutten-van Mölken MPMH, 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.

62. 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 Health Econ. 2017;18(6):671–83.

63. 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):239–51.

64. 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.

65. 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.

66. 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.

67. Gyrd-Hansen D. Is there additional value attached to health gains at the end of life? A revisit. Health Econ. 2018;27(1):e71–5.

68. 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.

69. Whitty JA, Lancsar E, Rixon K, Golenko X, Ratcliffe J. A Systematic Review of Stated Preference Studies Reporting Public Preferences for Healthcare Priority Setting. Patient. 2014;7(4):365–86.

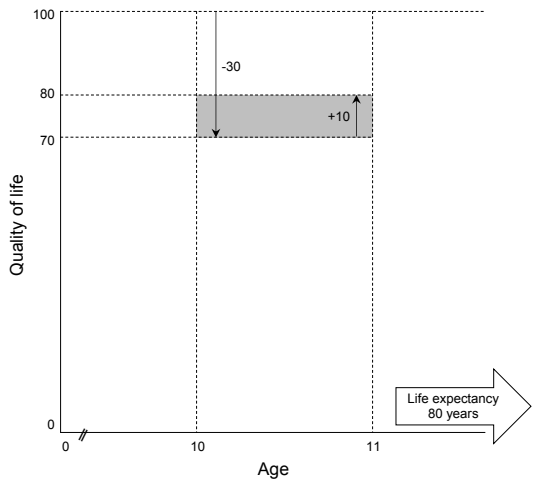
70. Richardson JRJ, McKie J, Peacock SJ, Iezzi A. Severity as an independent determinant of the social value of a health service. Eur J Health Econ. 2011;12(2):163–74.
71. van de Wetering EJ, van Exel NJA, Rose JM, Hoefman RJ, Brouwer WBF. Are some QALYs more equal than others? Eur J Health Econ. 2016;17(2):117–27.
72. van de Wetering L, van Exel J, Bobinac A, Brouwer WBF. Valuing QALYs in Relation to Equity Considerations Using a Discrete Choice Experiment. Pharmacoeconomics. 2015;33(12):1289–300.
73. National Institute for Health and Care Excellence. Modifiers: Task and finish group report. London; 2020.
74. National Institute for Health and Care Excellence. The NICE methods of health technology evaluation: The case for change. London; 2020.
75. 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):127–35.
76. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Biases in judgments reveal some heuristics of thinking under uncertainty. Science. 1974;185(4157):1124–31.
77. Rankin J, Robinson A. Accounting for protest zeros in contingent valuation: A review of literature. HEG Work Pap. 2018;(18–01).
78. 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 Health Econ. 2017;18(6):743–60.
79. Whynes DK, Sach TH. WTP and WTA: Do people think differently? Soc Sci Med. 2007;65(5):946–57.
80. Bobinac A. Mitigating hypothetical bias in willingness to pay studies: post-estimation uncertainty and anchoring on irrelevant information. Eur J Health Econ. 2019;20(1):75–82.
81. 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):337–61.
82. Johannesson M, Jönsson B. Economic evaluation in health care: Is there a role for cost-benefit analysis? Health Policy. 1991;17(1):1–23.
83. Shah KK, Tsuchiya A, Wailoo AJ. Valuing health at the end of life: An empirical study of public preferences. Eur J Health Econ. 2014;15(4):389–99.
84. 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.
85. 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.
86. 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.
87. Lancsar E, Gu Y, Gyrd-Hansen D, Butler J, Ratcliffe J, Bulfone L, et al. The relative value of different QALY types. J Health Econ. 2020;70:102303.

Appendix 4.1: Contingent-valuation task example (scenario 5)

Imagine a group of 10,000 patients who all have the same disease and to which you, your family, friends, and/or acquaintances can also belong. The patients would have lived in full health until the age of 80 had they not fallen ill.

The patients fall ill at the age of 10. The disease lasts for one year and leads to a lower quality of life during this year. Due to the disease, patients' quality of life decreases from 100 to 70 on the scale from 0 (dead) to 100 (full health).

A treatment is available that reduces the effects of the disease. The type and costs of the treatment are the same for all patients. Due to treatment, patients' quality of life during the year of illness is 10 points higher than without the treatment. Patients' quality of life will increase from 70 to 80 points. After this year, the patients fully recover and their quality of life returns to a score of 100.



To be able to pay for this treatment, the monthly basic health-insurance premium will increase for the duration of one year for all adult (18+) inhabitants of the Netherlands. Thus, this increase also applies to people who are not affected by the disease. After this year, the monthly premium will return to its current level. Without the increase in monthly premium, the patients will not be treated.

Please inspect the row with amounts below, from left to right, and state the highest basic health-insurance premium you are certainly willing to pay extra per month for the treatment that increases the 10-year-old patients' quality of life from 70 to 80 on the scale from 0 (dead) to 100 (full health).

Please be considerate of your net monthly household income when answering this question.

Please inspect the row with amounts below again, from left to right, and state the first basic health-insurance premium you are certainly not willing to pay extra per month for the treatment that increases the 10-year-old patients' quality of life from 70 to 80 on the scale from 0 (dead) to 100 (full health).

€0	€0.5	€1	€1.5	€2	€2.5	€3	€4	€5	€6	€7	€8	€10	€12	€14	€16	€18	€20	€22	€24	more
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Please be considerate of your net monthly household income when answering this question.

You have stated that you are certainly willing to pay €X extra basic health-insurance premium per month for the treatment that increases the 10-year-old patients' quality of life from 70 to 80 on the scale from 0 (dead) to 100 (full health), but certainly not more than €Y.

€0	€0.5	€1	€1.5	€2	€2.5	€3	€4	€5	€6	€7	€8	€10	€12	€14	€16	€18	€20	€22	€24	more
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

You have stated that you are certainly willing to pay €X extra basic health-insurance premium per month for the treatment that increases the 10-year-old patients' quality of life from 70 to 80 on the scale from 0 (dead) to 100 (full health), but certainly not more than €Y.

Within the range €X - €Y, what is the maximum amount you are willing to pay extra in monthly basic health-insurance premium for the treatment? Please be considerate of your net monthly household income when answering this question.

Appendix 4.2: Scenario characteristics

Table 4.2.1 Overview of scenario characteristics[a]

			Scenario characteristics					
Part	Module	Scenario	Age (at disease onset)	Severity[b]	QOL gain[c]	QOL after treatment[d]	Disease out-come[e]	Age (at death)
2		1	10	10	10	100	Full recovery	80
		2	20	10	10	100	Full recovery	80
		3	40	10	10	100	Full recovery	80
		4	70	10	10	100	Full recovery	80
		5	10	30	10	80	Full recovery	80
		6	20	30	10	80	Full recovery	80
		7	40	30	10	80	Full recovery	80
		8	70	30	10	80	Full recovery	80
		9	10	50	10	60	Full recovery	80
		10	20	50	10	60	Full recovery	80
		11	40	50	10	60	Full recovery	80
		12	70	50	10	60	Full recovery	80
		13	10	70	10	40	Full recovery	80
		14	20	70	10	40	Full recovery	80
		15	40	70	10	40	Full recovery	80
		16	70	70	10	40	Full recovery	80
		17	10	90	10	20	Full recovery	80
		18	20	90	10	20	Full recovery	80
		19	40	90	10	20	Full recovery	80
		20	70	90	10	20	Full recovery	80
3	1	21	10	50	20	70	Full recovery	80
		22	20	50	20	70	Full recovery	80
		23	40	50	20	70	Full recovery	80
		24	70	50	20	70	Full recovery	80
	2	25	10	50	50	100	Full recovery	80
		26	20	50	50	100	Full recovery	80
		27	40	50	50	100	Full recovery	80
		28	70	50	50	100	Full recovery	80
	3	29	10	50	20	70	Death	11
		30	20	50	20	70	Death	21
		31	40	50	20	70	Death	41
		32	70	50	20	70	Death	71
	4	33	10	50	50	100	Death	11
		34	20	50	50	100	Death	21
		35	40	50	50	100	Death	41
		36	70	50	50	100	Death	71

QOL, health-related quality of life (measured in points on a visual analogue scale (VAS) ranging from 0 “death” to 100 “full health”); [a] The number of patients is 10,000 in each scenario; [b] Severity is operationalised in terms of disease-related QOL loss and measured in points from 100 on the VAS; [c] Treatment-related QOL gain is measured in points on the VAS; [d] QOL after treatment is measured in points on the VAS and calculated as: 100 – disease-related QOL loss + treatment-related QOL gain; [e] Disease outcome is one year after falling ill (patients who fully recover return to having a QOL of 100 on the VAS).

Chapter 5
Willingness to pay for quality and length of life gains in end of life patients of different ages

Based on: Reckers-Droog VT, van Exel NJA, Brouwer WBF. Willingness to pay for quality and length of life gains in end of life patients of different ages. Submitted. 2021.

Abstract

Health gains are increasingly weighted in economic evaluations of new health technologies to guide resource-allocation decisions in healthcare. In England, for example, a higher weight is attached to quality-adjusted life-years (QALYs) gained from life-extending end-of-life (EOL) treatments. Societal preferences for QALY gains in EOL patients are increasingly examined. Although the available evidence suggests that gains in health-related quality of life (QOL) may be preferred to gains in life expectancy (LE), little is known about the influence of EOL patients' age on these preferences. In this study, we examine the willingness to pay (WTP) for QOL and LE gains in EOL patients of different ages in a sample (n=803) of the general public in the Netherlands. We found that WTP was relatively higher for QOL and LE gains in younger EOL patients. We further found indications suggesting that WTP may be relatively higher for QOL gains at the EOL. However, relative preferences for QOL and LE gains appeared to be independent of the age of EOL patients. Our results may inform discussions on attaching differential weights to QOL and LE gains in EOL patients of different ages with the objective to better align resource-allocation decisions with societal preferences.

5.1 Introduction

Health gains are increasingly weighted in economic evaluations of new health technologies to guide healthcare priority setting. Under strict conditions, the National Institute for Health and Care Excellence (NICE) in England, attaches a higher weight to quality-adjusted life-year (QALY) gains in patients at the end of life (EOL) [1]. The weight can be applied to QALYs gained from life-extending end-of-life treatments if, and only if, the following conditions are met: (i) the health technology under evaluation is indicated for patients with a short life expectancy, normally of less than 24 months, (ii) there is sufficient evidence that the technology prolongs patients' lives by at least three months, as compared to current treatment, and (iii) the technology is licensed or otherwise indicated for small patient populations. The maximum weight that can be applied ranges between 1.7 and 2.5, so that the adjusted incremental cost-effectiveness ratio (ICER) of the technology can be evaluated against the normal threshold range of £20,000 to £30,000 per QALY gained [1]. Note that this equals a maximum threshold of £50,000 for an unadjusted ICER.

Since NICE introduced this 'EOL premium' in 2009 [1], societal preferences for health gains in EOL patients (as compared to non-EOL patients) are increasingly examined [2]. Possibly because the premium solely applies to life-extending EOL treatments [1], most of these studies examined societal preferences solely in relation to gains in life expectancy (LE) [2]. Only a few studies (also) examined these preferences in relation to gains in health-related quality of life (QOL). The results of studies that examined preferences for both types of health gains seem to suggest that societal preferences for QOL gains may be stronger than for LE gains in EOL patients, thus calling into question NICE's EOL premium [e.g. 3,4]. The extent to which these preferences are influenced by the age of EOL patients remains largely unexplored [5].

Based on empirical evidence that societal preferences for health gains in younger patients may be relatively strong [6,7], it seems plausible to hypothesise that societal preferences for QOL and LE gains in younger EOL patients would be stronger than for similar gains in older patients. Based on empirical evidence that societal preferences for health gains and losses may be dependent on what is considered an acceptable health state at different ages [8,9], it also seems plausible to hypothesise that the relative preferences for QOL and LE gains would be different for EOL patients of different ages. As there is little evidence on the role of age in this context, we explore whether there is indeed any evidence of this by examining the relative strength of societal preferences for QOL and LE gains in EOL patients of different ages.

We elicited preferences in terms of the willingness to pay (WTP) for the health gains and, given the aim of our study, focused on the relative rather than absolute height of the WTP. Our results may inform discussions in countries that consider attaching differential weights to QALYs gained from life-exten-

ding EOL treatments, like England, or to QOL and LE gains in EOL patients of different ages with the objective to (better) align resource-allocation decisions with societal preferences.

5.2 Methods

5.2.1 Sample and data collection

We collected data as part of a larger contingent-valuation (CV) study in August 2019 [10]. Here, we report on the WTP elicited from 803 respondents who were quota sampled to be representative of the general public in the Netherlands by age (18-75 years), sex, and education level and to cover a broad range of household incomes.

We asked respondents to complete the tasks from a socially-inclusive-personal (SIP) perspective and take into consideration that they themselves, their family, friends, and/or acquaintances as well as unknown others could be part of the hypothetical patient group. This perspective represents a combination of the personal and social perspectives [11] and facilitated the use of an increase in monthly health-insurance premium as payment vehicle, which is how adult inhabitants of the Netherlands contribute to the collectively funded healthcare system.

5.2.2 Questionnaire

At the start of the questionnaire (see Figure 5.1), we introduced respondents to the following concepts: (i) QOL, operationalised in points on a visual analogue scale (VAS) ranging from 0 "dead" to 100 "full health", (ii) severity, operationalised in terms of disease-related QOL loss (in points from 100 on the VAS) and LE loss (in years from 80 life years), and (iii) treatment-related QOL and LE gains. We familiarised respondents with the concepts and tasks by asking them to assess their own QOL "today" on the VAS and complete one practice task from a personal perspective. Respondents then assessed the level of clarity of the practice task on a seven-point Likert scale ranging from 1 "very unclear" to 7 "very clear" and indicated on what expenses they would likely cut back to cover the stated WTP. We asked the latter to increase respondents' awareness of the associated opportunity costs.

We then asked respondents to complete three tasks from a SIP perspective. For the first two tasks, we randomly assigned respondents to two out of 20 scenarios in which patients fully recovered one year after falling ill. For the third task, we randomly assigned respondents to one out of eight scenarios in which patients died one year after falling ill. These eight scenarios were evenly distributed across two modules and are the main focus of this study (see Appendix 5.1 and 5.2 for an overview of the characteristics and a description of the scenarios). Each of the scenarios described a group of 10,000 patients

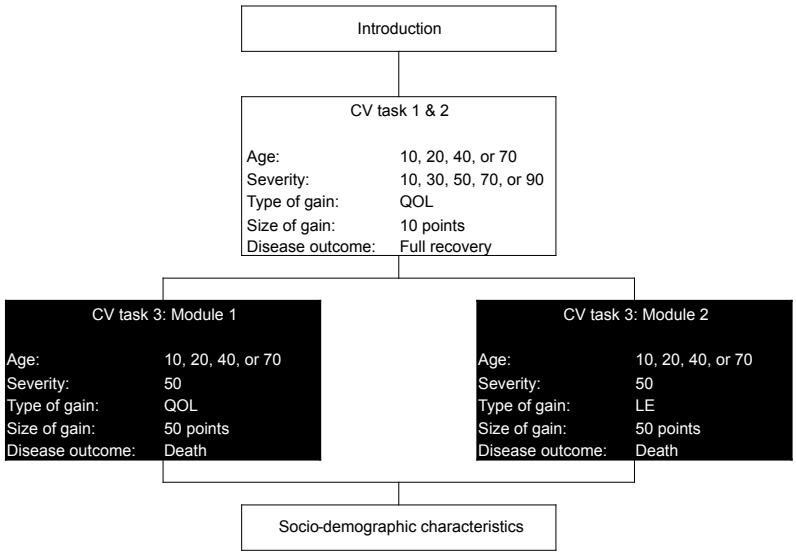


Fig. 5.1 Questionnaire structure[a]
CV, contingent-valuation; LE, life expectancy; QOL, health-related quality of life (measured in points on a visual analogue scale, ranging from 0 "dead" to 100 "full health"); [a] CV task 3 (i.e. modules 1 and 2) is the main focus of the current study and, therefore, presented in contrasting black (note that the questionnaire included three additional modules on which we report elsewhere [10]; Severity is operationalised in terms of disease-related QOL loss and measured in points from 100 on the VAS; Disease outcomes "Full recovery" and "Death" occur one year after falling ill.

who would have lived in full health (a score of 100 on the VAS) until the age of 80 had they not fallen ill at age 10, 20, 40, or 70. Due to the disease, patients' QOL decreased from 100 to 50 on the VAS for the duration of one year, after which they died. We explained that a treatment was available that improved patients' health with 50 points.

In module 1, treatment increased patients' QOL from 50 to 100 points on the VAS but did not affect their LE (this remained one year). In module 2, treatment increased patients' LE with one additional year but did not affect their QOL (this remained 50 on the VAS). We explained that the treatment could be made available to patients by increasing the monthly health-insurance premium for all adult inhabitants of the Netherlands for the duration of one year. We elicited respondents' WTP for the health gains by applying a two-step CV procedure (see Appendix 5.2 for a description of the procedure and a task example). At the end of the questionnaire, we asked respondents about their socio-demographic characteristics.

5.2.3 Statistical analyses

Before conducting the analyses, we classified WTPs of €0 as either a true or protest zero valuation (see Appendix 5.2 and 5.3 for the classification criteria and distribution of true and protest zero valuations across scenarios). We excluded protest zero valuations, outlying WTPs (z-score ≥1.64, determined based on the distribution of raw WTPs), and speeders (minimum completion time of 90 seconds, determined based on a timed test by three researchers).

We first analysed the data obtained from the third task. We calculated the mean (SD; 95% CI) and differences in mean (SE; 95% CI) WTP for QOL and LE gains in EOL patients of different ages (see Appendix 5.3 for median (IQR) WTPs). We applied two-tailed Welch's t-tests (Bonferroni corrected) to examine whether WTP for QOL and LE gains was *different at different ages*. We bootstrapped (5000 repetitions) the differences in means to examine whether WTP for LE gains and QOL gains was relatively different *between* different ages. We further applied Ordinary Least Squares (OLS) models to examine the (interaction) effect of patients' age and the type of health gain on WTP.

We then analysed the data obtained from all three tasks to account for the panel structure of the data and allow for individual effects. We applied random-effects Generalised Least Squares (GLS) models to further examine the (interaction) effect of patients' age, the type of health gain, and respondent characteristics on WTP. We controlled for the scenario characteristics of the first two tasks, i.e. severity, age, size of the health gain, disease outcome (the latter two variables coded as 0 "10 points/full recovery after one year" and 1 "50 points/death after one year" to avoid perfect multicollinearity), and a time-effect (labelled "CV task") in all GLS models.

We performed sensitivity analyses to assess the robustness of our results by regressing log(WTP+1) on scenario and respondent characteristics and by examining the effect of respondents' proximity to the age of patients (coded as |Age respondents – Age patients|) on WTP. We assessed the latter as respondents closer to the specified age of patients may have had strong(er) self-regarding preferences, and hence a higher WTP in some scenarios. We further assessed robustness by examining the effect of the WTP stated in the practice task and by repeating the analyses excluding respondents with a low clarity score (i.e. 1–3 level) for the practice task and relaxing the minimum completion time to 39 seconds (z-score ≤-1.64, based on the distribution of completion times).

We conducted the analyses using Stata 16.1 (Stata Corp LP, College station, Texas).

5.3 Results

Table 5.1 presents the descriptive statistics of the sample (n=495) that remained after excluding protest zero valuations (n=86), outliers (n=25), and speeders (n=197).

Table 5.1 Sample characteristics (n=495)[a]

	%	Mean (SD)
Age (Years)		52.4 (16.3)
Sex (Female)	49.3	
Education level[b]		
Low	12.7	
Medium	55.6	
High	31.7	
Household income (After tax)		
<€1,999	32.3	
€2,000 – €3,999	39.8	
≥€4,000	22.4	
NS	5.5	
Children (Yes)	59.8	
QOL (0-100 VAS)		80.5 (16.9)
Completion time of CV tasks (Minutes)		5.1 (8.4)

CV, contingent-valuation; NA, Not Applicable; NS, Not Stated; QOL, health-related quality of life (measured in points on a visual analogue scale, ranging from 0 "dead" to 100 "full health"); [a] Respondents who gave protest zero valuations (n=86) or stated an outlying raw WTP (n=25) in the third CV task as well as respondents who completed the three CV tasks in less than 90 seconds (n=197) are excluded from this table. Of the respondents who completed the tasks too quickly, 10 also stated an outlying WTP; [b] Low = lower vocational and primary school, Medium = middle vocational and secondary school, High = higher vocational and academic education.

Table 5.2 presents the mean (SD; 95% CI) WTP for QOL and LE gains in EOL patients of different ages and the difference in means (SE; 95% CI). On average, WTP was €8.7 per month for the duration of one year for QOL gains and €7.9 for LE gains. WTP was higher for QOL and LE gains in younger than in older EOL patients, except for a relatively low WTP for QOL gains in EOL patients aged 20. The results indicate that WTP for QOL and LE gains were similar in EOL patients aged 10, but higher for LE than for QOL gains in EOL patients aged 20. WTP was higher for QOL than for LE gains for EOL patients aged 40 and 70. The observed differences were not statistically significantly different from 0 at different ages, nor between different ages ($p>0.10$).

Table 5.3 presents the OLS regression results. These results indicate that compared to EOL patients aged 10, a higher age was, ceteris paribus, associated with a lower WTP for health gains (model 1: β -1.48 to -2.60). The difference in WTP for QOL and LE gains was not statistically significant, neither was the interaction between patients' age and the type of health gain.

Table 5.2 Mean (SD; 95% CI) WTP for QOL and LE gains in end-of-life patients of different ages (n=495)[a]

	QOL gain[b]			LE gain[c]			Difference (QOL gain – LE gain)	
Age	n	Mean (SD)	95% CI	n	Mean (SD)	95% CI	ΔMean (SE)	95% CI
10	71	9.8 (6.7)	8.2, 11.4	51	9.8 (9.5)	7.1, 12.5	0.0 (1.5)	-3.1, 2.9
20	64	7.8 (6.2)	6.3, 9.4	62	8.7 (6.9)	7.0, 10.5	-0.9 (1.2)	-3.2, 1.4
40	61	8.9 (8.1)	6.8, 10.9	62	7.2 (6.0)	5.7, 8.7	1.7 (1.3)	-0.8, 4.2
70	60	8.0 (6.7)	6.2, 9.8	64	6.3 (6.4)	4.7, 7.9	1.7 (1.2)	-0.7, 4.0
n	256			239				
Average		8.7 (7.0)	7.8, 9.5		7.9 (7.3)	7.0, 8.8	0.8 (0.6)	-0.5, 2.0

LE, life expectancy; QOL, health-related quality of life (measured in points on a visual analogue scale (VAS), ranging from 0 “dead” to 100 “full health”); WTP, willingness to pay (in € per month for the duration of one year); [a] Respondents with protest zero valuations, outlying WTPs (Module 1: ≥€32.80; Module 2: ≥€46.75), and those who completed the tasks in less than 90 seconds are excluded from this table; [b] Module 1: Treatment-related QOL gain is 50 points, and hence patients’ QOL is restored to 100 points on the VAS for the duration of one year, after which they will die; [c] Module 2: Treatment-related LE gain is 1 year, and hence patients’ will live one additional year with a QOL of 50 points on the VAS, after which they die.

Table 5.3 OLS regression results

	Model 1		Model 2	
DV: WTP	β (SE)	95% CI	β (SE)	95% CI
Age (10 = reference)	-	-	-	-
20	-1.48* (0.90)	-3.25, 0.28	-1.98 (1.22)	-4.38, 0.42
40	-1.73* (0.91)	-3.51, 0.05	-0.95 (1.24)	-3.38, 1.48
70	-2.60*** (0.91)	-4.38, -0.83	-1.80 (1.24)	-4.24, 0.64
LE gain (QOL gain = reference)	-0.62 (0.64)	-1.87, 0.64	-0.02 (1.30)	-2.57, 2.53
Age*LE gain (Age 10; QOL gain = reference)			-	-
20*LE gain			0.92 (1.81)	-2.64, 4.48
40*LE gain			-1.66 (1.82)	-5.24, 1.92
70*LE gain			-1.67 (1.82)	-5.25, 1.90
Constant	10.06*** (0.69)	8.70, 11.43	9.81*** (0.84)	8.16, 11.46
R^2	0.02		0.03	
n	495			

DV, dependent variable; * p<0.10, ** p<0.05, *** p<0.01.

Table 5.4 on the next page presents the GLS regression results. These results confirm that compared to patients aged 10, a higher age was, ceteris paribus, associated with a lower WTP for health gains (model 3 to 5: β -0.56 to -1.00). Note that the difference in WTP for health gains in patients aged 20, 40, and 70 became smaller when controlling for the scenario characteristics of CV tasks 1 and 2, and for respondent characteristics. Compared to QOL gains, LE gains were, ceteris paribus, associated with a lower WTP. However, this estimate was only statistically significant when modelled as a main effect (model 3: β -1.28), not when modelled as a conditional effect (models 4 and 5: β -1.39 and -1.22). The interaction between patients' age and the type of health gain indicates that the WTP for QOL and LE gains was independent of patients' age. The results further indicate that WTP generally increased with increased severity (models 3 to 5: β 0.40 to 1.76) and was higher for a health gain of 50 points in patients who died than for a health gain of 10 points in patients who fully recovered one year after falling ill (models 3 to 5: β 1.99 to 2.08). Note that it was not possible to distinguish between the effect of the size of the health gain and patients' disease outcome on WTP. A higher respondent age was further associated with a lower WTP (model 5: β -0.28), whereas having children (model 5: β 1.65) and a higher household income (model 5: β 1.78) were associated with a higher WTP. The sensitivity analyses indicated that the WTP stated in the practice task (see Appendix 5.4, model 5B: β 0.01, $p<0.001$) had a marginal effect on WTP and that our results were robust.

5.4 Discussion

This study was one of the first to examine societal preferences for QOL and LE gains in EOL patients of different ages. We found that WTP is relatively higher for QOL and LE gains in younger patients, which might be explained by preferences relating to lifetime health [8]. We further found indications of a higher WTP for QOL gains than for LE gains at the EOL; however, relative preferences for QOL and LE gains appeared to be independent of the age of EOL patients.

Reckers-Droog et al. [10] discuss the main strengths and limitations of our design. Here, we want to highlight that we applied a SIP perspective for eliciting WTP, and hence cannot clearly distinguish between (unobservable) self-regarding and other-regarding preferences. Nonetheless, this perspective likely aligns most with actual decisions regarding higher payments to a collectively funded healthcare system like that of the Netherlands. We also want to highlight that we elicited preferences under certainty (in terms of QALYs gained) in hypothetical scenarios, in which we standardised the healthy-life expectancy and total health gain of the patient groups. This aimed to increase the clarity of the tasks and reduce the possible influence of other considerations (e.g. associated with health maximisation). Nonetheless, this came at the expense of realism and generalisability, and may have influenced our results. For example, WTP may have been different if the scenarios described patients

Table 5.4 GLS regression results

	Model 3		Model 4		Model 5	
DV: WTP	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
Age (10 = reference)	-	-	-	-	-	-
20	-0.57* (0.34)	-1.24, 0.10	-0.75** (0.34)	-1.41, -0.09	-0.86** (0.36)	-1.60, -0.16
40	-0.56* (0.31)	-1.16, 0.04	-0.56* (0.31)	-1.17, 0.04	-0.56* (0.32)	-1.19, 0.08
70	-1.00*** (0.34)	-1.65, -0.34	-0.89*** (0.33)	-1.53, -0.24	-0.86** (0.34)	-1.54, -0.19
LE gain (QOL gain = reference)	-1.28*** (0.45)	-2.17, -0.39	-1.39 (1.14)	-3.61, 0.84	-1.22 (1.20)	-3.58, 1.14
Age*LE gain (Age 10; QOL gain = reference)			-	-	-	-
20*LE gain			1.07 (1.22)	-1.32, 3.46	1.08 (1.30)	-1.47, 3.63
40*LE gain			0.00 (1.30)	-2.56, 2.55	-0.27 (1.38)	-2.98, 2.44
70*LE gain			-0.63 (1.22)	-3.02, 1.76	-0.86 (1.28)	-3.37, 1.66
Severity[a] (10 = reference)	-	-	-	-	-	-
30	0.40 (0.45)	-0.49, 1.29	0.43 (0.45)	-0.45, 1.30	0.34 (0.45)	-0.54, 1.23
50	0.92** (0.41)	0.13, 1.72	0.94** (0.40)	0.15, 1.73	0.90** (0.41)	0.08, 1.71
70	0.91** (0.39)	0.15, 1.67	0.91** (0.39)	0.15, 1.67	0.92**(0.40)	0.14, 1.69
90	1.73*** (0.43)	0.88, 2.58	1.71*** (0.44)	0.86, 2.57	1.76*** (0.45)	0.88, 2.64
Health gain 50 points + death (Health gain 10 points + full recovery = reference)	1.99*** (0.43)	1.14, 2.84	2.00*** (0.43)	1.15, 2.85	2.08*** (0.45)	1.19, 2.97
Age (of respondents)					-0.28*** (0.11)	-0.50, -0.06
Age^2 (of respondents)					0.00*** (0.00)	0.00, 0.00
Sex (Female = reference)					0.80 (0.59)	-0.36, 1.96
Children (No = reference)					1.65* (0.64)	0.39, 2.91
Education level[b] (Low = reference)					-	-
Medium					-0.28 (1.12)	-2.48, 1.92
High					-0.77 (1.30)	-3.32, 1.79
Household income[c] (Thousands)					1.78*** (0.43)	0.95, 2.62
QOL (0-100 VAS)					-0.01 (0.02)	-0.04, 0.03
CV task	-0.79*** (0.19)	-1.16, -0.42	-0.80*** (0.19)	-1.17, -0.43	-0.79*** (0.20)	-1.17, -0.41
Constant	8.91*** (0.50)	7.93, 9.89	8.93*** (0.49)	7.97, 9.88	11.52*** (2.94)	5.76, 17.29
R^2 overall	0.01		0.01		0.08	
n observations; groups	1,438; 495		1,438; 495		1,360; 468	

CV, contingent valuation; DV, dependent variable; LE, life expectancy; NS, Not Stated; QOL, health-related quality of life (measured in points on a VAS); VAS, visual analogue scale (ranging from 0 “dead” to 100 “full health”); WTP, willingness to pay (in € per month for the duration of one year); a Severity 50 coefficients represent the impact on WTP of disease severity level 50 of patients who fully recovered or who died one year after falling ill; b Low = lower vocational and primary school, Medium = middle vocational and secondary school, High = higher vocational and academic education; c Household income is adjusted for household size and calculated as household income/(household size)^0.5 to account for economies of scale [12]. In model 5, respondents who did not state their household income (n=27) are excluded, and hence the number of observations and groups is lower than in models 4 and 5; * p<0.10, ** p<0.05, *** p<0.01.

at risk of falling ill or if treatment improved patients' health with a specific probability. In addition, two limitations that are specific to the current study deserve emphasis. First, the QOL gains of 50 points in module 1 fully restored patients' QOL to 100 points in their last year of life, which means that we cannot distinguish between the effect of QOL gains and patients' health being fully restored on WTP. Second, we did not separate the effect of patients' age in the first two tasks (where patients gained 10 points and fully recovered) from the effect of their age in the third task (where patients gained 50 points and died) in the GLS models. However, this enabled us to assess the effect of patients' age on WTP without introducing multicollinearity. Note that the direction and relative strength of the age coefficients correspond between the OLS and GLS models, indicating that our conclusions regarding age-related preferences in society may be robust and independent of the size of health gains and disease outcome of patients.

With this study, we aimed to contribute to the limited evidence on societal preferences for QOL and LE gains in EOL patients of different ages. Our findings confirm previous findings that societal preferences for health gains in younger patients are relatively strong and extent these findings to the EOL context. They also confirm previous findings that dispute NICE's application of a premium solely for LE gains, rather than for QOL gains or a combination of both, at the EOL. As such, they may inform discussions in countries that consider attaching differential weights to QOL and LE gains in EOL patients of different ages, aiming to better align resource-allocation decisions with societal preferences.

References

1. National Institute for Health and Care Excellence. Appraising life-extending, end of life treatments. MeReC Bulletin. London; 2009.
2. 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.
3. 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.
4. 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.
5. Lancsar E, Gu Y, Gyrd-Hansen D, Butler J, Ratcliffe J, Bulfone L, et al. The relative value of different QALY types. J Health Econ. 2020;70:102303.
6. Bobinac A, van de Wetering L, van Exel NJA, Brouwer WBF. Equity-dependent social willingness to pay for a QALY. In: Valuing and refining outcome measures for economic evaluations in health care. 2015. p. 83–108.
7. Reckers-Droog VT, van Exel NJA, Brouwer WBF. Equity Weights for Priority Setting in Healthcare: Severity, Age, or Both? Value Health. 2019;22(12):1441–9.
8. Olsen JA. Priority preferences: "End of life" does not matter, but total life does. Value Health. 2013;16(6):1063–6.
9. 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.
10. Reckers-Droog VT, van Exel NJA, Brouwer WBF. Willingness to pay for health-related quality of life gains in relation to the disease severity and age of patients. Forthcoming Value Health. 2021.
11. 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):545–51.
12. Johnson DS, Smeeding TM, Torrey BB. Economic inequality through the prisms of income and consumption. Mon Labor Rev. 2005;128(4):11–24.
13. Ministry of Health, Welfare and Sport. Wat kost het om een zorgverzekering af te sluiten? [Internet]. [cited 2020 Jan 13]. Available from: https://www.rijksoverheid.nl/onderwerpen/zorgverzekering/vraag-en-antwoord/premie-zorgverzekering

Appendix 5.1: Scenario characteristics

Table 5.1.1 Overview of scenario characteristics[a]

Module	Scenario	Scenario characteristics					
		Age (at disease onset)	Severity[b]	QOL gain[c]	QOL after treatment[d]	Disease outcome[e]	Age (at death)
	1	10	10	10	100	Full recovery	80
	2	20	10	10	100	Full recovery	80
	3	40	10	10	100	Full recovery	80
	4	70	10	10	100	Full recovery	80
	5	10	30	10	80	Full recovery	80
	6	20	30	10	80	Full recovery	80
	7	40	30	10	80	Full recovery	80
	8	70	30	10	80	Full recovery	80
	9	10	50	10	60	Full recovery	80
	10	20	50	10	60	Full recovery	80
	11	40	50	10	60	Full recovery	80
	12	70	50	10	60	Full recovery	80
	13	10	70	10	40	Full recovery	80
	14	20	70	10	40	Full recovery	80
	15	40	70	10	40	Full recovery	80
	16	70	70	10	40	Full recovery	80
	17	10	90	10	20	Full recovery	80
	18	20	90	10	20	Full recovery	80
	19	40	90	10	20	Full recovery	80
	20	70	90	10	20	Full recovery	80
1	21	10	50	50	100	Death	11
	22	20	50	50	100	Death	21
	23	40	50	50	100	Death	41
	24	70	50	50	100	Death	71
2	25	10	50	50	50	Death	12
	26	20	50	50	50	Death	22
	27	40	50	50	50	Death	42
	28	70	50	50	50	Death	72

QOL, health-related quality of life (measured in points on a visual analogue scale (VAS) ranging from 0 “death” to 100 “full health”); [a] The number of patients is 10,000 in each scenario; [b] Severity is operationalised in terms of disease-related QOL loss and measured in points from 100 on the VAS; [c] Treatment-related QOL gain is measured in points on the VAS; [d] QOL after treatment is measured in points on the VAS and calculated as: 100 – disease-related QOL loss + treatment-related QOL gain; [e] Disease outcome is one year after falling ill (patients who fully recover return to having a QOL of 100 on the VAS).

Appendix 5.2: Applied perspectives and contingent-valuation procedure

Applied perspectives

1. Practice task: individual perspective

At the start of the scenario, respondents were asked to imagine that their quality-of-life (QOL) was 100 on a visual analogue scale (VAS) ranging from 0 "dead" to 100 "full health" and that—due to some disease—their QOL decreased to 40 on the VAS. The disease would last for one year and after this year they would return to full health (a score of 100 on the VAS). During this year, respondents' QOL could be improved from 40 to 80 points on the VAS by taking a medicine that would not have any negative side effects. This medicine would not be covered by their health-insurance and would have to be paid out of pocket on a monthly basis. We elicited respondents' willingness to pay (WTP) for the treatment-related QOL gain by applying the two-step procedure described below.

2. Task 1 and 2: socially-inclusive-personal perspective

For task 1 and 2, respondents were randomly assigned to two 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 one year. After this year, they would fully recover (i.e. return to a score of 100 on the VAS). The disease would not affect patients' life expectancy (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 one 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 two-step procedure described below.

3. Task 3: socially-inclusive-personal perspective

For task 3, respondents were randomly assigned to one out of eight scenarios that were evenly distributed across two modules. 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, patients' QoL decreased from 100 to 50 on the VAS for the duration of one

year, after which the patients would die. We explained that a treatment was available that improved patients' health by 50 points and that the treatment type and costs were the same for all patients. In module 1, the treatment increased patients' QoL from 50 to 100 on the VAS, while it did not affect their LE (this remained one year). In module 2, the treatment did not affect patients' QoL (this remained 50), but it increased their LE with one additional year. The treatment could be made available to patients by increasing the monthly basic health-insurance premium for the duration of one 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 two-step procedure described below.

Contingent-valuation procedure

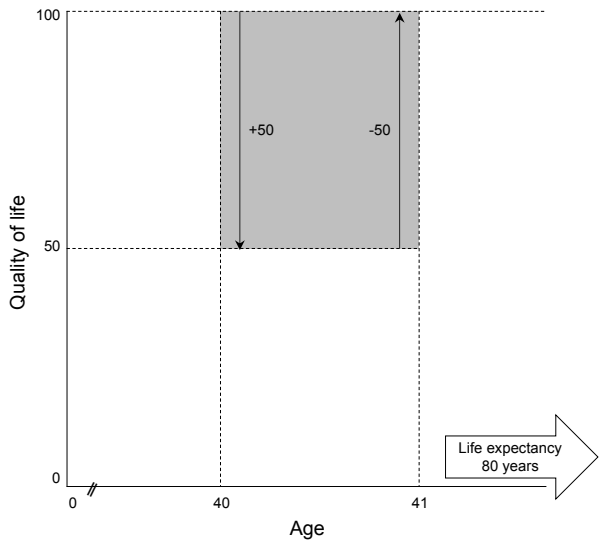
We elicited the willingness to pay (WTP) for the health gains by applying a two-step contingent-valuation procedure. In step one, we presented a payment scale that ranged from a €0 to €24 increase in monthly health-insurance premium with unevenly distributed intervals between the value points (i.e. €0, €0.50, €1, €1.50, €2, €2.50, €3, €4, €5, €6, €7, €8, €10, €12, €14, €16, €18, €20, €22, €24, and "more"). 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 one 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 one year. We asked respondents who stated a WTP of €0 to explain their main reason for having this preference by completing an open-text field or by checking one of six randomised answer options that either related to a true zero valuation or a protest zero valuation. The open-text field answers were qualitatively assessed by the first two authors and subsequently classified as either true or protest zero valuation. The answer options relating to true zero valuations were "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". The answer options relating to protest zero valuations were "I am against an increase in monthly health-insurance premium", "Patients should pay for the treatment themselves", and "The value of health and healthcare cannot be expressed in monetary terms". In step two, we asked respondents to indicate the maximum increase in monthly health-insurance premium they would be willing to pay within the payment range obtained in step one. In both steps, we asked respondents to take their net monthly household income (as a proxy for their ability to pay) and the contingency that the stated WTP would be mandatory for all adult inhabitants of the Netherlands into account. Note that, by approximation, the number of adults is 13.7 million and the monthly health-insurance premium is €115.00 per person in 2019 [13].

Example of a contingent-valuation task (Module 1, scenario 23)

Imagine a group of 10,000 patients who all have the same disease and to which you, your family, friends, and acquaintances can also belong. The patients would have lived in full health (a score of 100) until the age of 80 had they not fallen ill.

The patients fall ill at the age of 40. Due to the disease, patients' quality of life decreases from 100 to 50 for the duration of one year. After this year the patients will die.

A treatment is available that reduces the effects of the disease. The treatment and the costs of treatment are the same for all patients. Due to treatment, patients' quality of life will be 50 points higher during their last year of life than without the treatment. Hence, patients' quality of life will increase from 50 to 100 points. After this year, the patients will die.



To be able to pay for this treatment, the monthly health-insurance premium will increase for the duration of one year for all adult (18+) inhabitants of the Netherlands. Thus, this increase also applies to people who are not affected by the disease. After this year, the monthly premium will return to its current level. Without the increase in monthly premium, the patients will not receive the treatment.

Please inspect the row with euro amounts below, from left to right, and indicate the highest health-insurance premium you are certainly willing to pay extra per month for the treatment that increases the 40-year-old patients'

quality of life from 50 to 100 on the scale from 0 (dead) to 100 (full health) during their last year of life.

Please be considerate of your net monthly household income when answering this question.

€0	€0.5	€1	€1.5	€2	€2.5	€3	€4	€5	€6	€7	€8	€10	€12	€14	€16	€18	€20	€22	€24	more
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Please inspect the row with euro amounts below again, from left to right, and indicate the first health-insurance premium you are certainly not willing to pay extra per month for the treatment that increases the 40-year-old patients' quality of life from 50 to 100 on the scale from 0 (dead) to 100 (full health) during their last year of life.

Please be considerate of your net monthly household income when answering this question.

€0	€0.5	€1	€1.5	€2	€2.5	€3	€4	€5	€6	€7	€8	€10	€12	€14	€16	€18	€20	€22	€24	more
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

You have indicated that you are certainly willing to pay €X extra health-insurance premium per month for the treatment that increases the 40-year-old patients' quality of life from 50 to 100 on the scale from 0 (dead) to 100 (full health) during their last year of life., but certainly not more than €Y.

Within the range €X - €Y, what is the maximum amount you are willing to pay extra in monthly health-insurance premium for this treatment? Please be considerate of your net monthly household income when answering this question.

Appendix 5.3: Zero valuations and median willingness to pay

Table 5.3.1 Distribution of true and protest zero valuations across scenarios[a]

QOL gain[b]			LE gain[c]		
Scenario[d]	True zero (n)	Protest zero (n)	Scenario[d]	True zero (n)	Protest zero (n)
21	1	11	25	7	15
22	7	9	26	4	9
23	5	13	27	7	12
24	1	7	28	8	10
n	14	40	n	26	46

LE, life expectancy; QOL, health-related quality of life (measured on a visual analogue scale (VAS), ranging from 0 "dead" to 100 "full health"); [a] See Appendix B for the classification criteria. [b] Treatment-related QOL gain is 50 points, and hence patients' QOL is restored to 100 points on the VAS for the duration of one year, after which the patients die. [c] Treatment-related LE gain is 1 year, and hence patients' will live one additional year (i.e. two years from onset of the disease) with a QOL of 50 points on the VAS, after which they will die. [d] See Appendix 5.1 for an overview of scenario characteristics.

Table 5.3.2 Median (IQR) WTP for QOL and LE gains in end-of-life patients of different ages (n = 520)[a]

	QOL gain[b]		LE gain[c]		Difference (QOL gain – LE gain)
Age	n	Median (IQR)	n	Median (IQR)	ΔMedian (IQR)
10	74	10.0 (4.0 – 15.0)	55	9.0 (3.0 – 20.0)	1.0 (1.0 – -5.0)
20	67	7.0 (2.5 – 14.0)	65	8.0 (2.8 – 14.0)	-1.0 (-0.3 – 0.0)
40	63	6.0 (2.5 – 15.0)	65	6.0 (2.5 – 12.0)	0.0 (0.0 – 3.0)
70	66	7.3 (2.5 – 15.0)	65	4.0 (2.0 – 10.0)	3.3 (0.5 – 5.0)
n	270		250		
Average		7.5 (3.0 – 15.0)		6.5 (2.5 – 12.0)	1.0 (0.5 – 3.0)

IQR, interquartile range; LE, life expectancy; QOL, health-related quality of life (measured in points on a visual analogue scale (VAS), ranging from 0 "dead" to 100 "full health"); WTP, willingness to pay (in € per month for the duration of one year); [a] Respondents with protest zero valuations (n=86) and those who completed the tasks in less than 90 seconds (n=197) are excluded from this table. In the scenarios, the groups consist of 10,000 patients with a severity level of 50 points on the VAS; [b] Treatment-related QOL gain is 50 points, and hence patients' QOL is restored to 100 points on the VAS for the duration of one year, after which the patients die; [c] Treatment-related LE gain is 1 year, and hence patients' will live one additional year (i.e. two years from onset of the disease) with a QOL of 50 points on the VAS, after which they will die.

Appendix 5.4: Sensitivity analyses

Table 5.4.1 Results sensitivity analyses (based on Table 5.4, model 5)[a]

	Model 5[A]		Model 5[B]		Model 5[C]		Model 5[D]	
	Assess influence proximity of respondents' age to that of patients		Corrected for WTP stated in practice task		Excluding respondents with low clarity score (1−3) for practice task		Minimum time to complete 39 seconds	
DV: WTP	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
Age[b] (10 = reference)			-	-	-	-	-	-
20	-	-	-0.88** (0.37)	-1.59, -0.16	-0.86** (0.36)	-1.56, -0.16	-0.80** (0.32)	-1.41, -0.18
40			-0.55* (0.33)	-1.20, 0.09	-0.56* (0.32)	-1.19, 0.08	-0.52** (0.30)	-1.10, -0.07
70			-0.88** (0.35)	-1.57, -0.20	-0.86** (0.34)	-1.54, -0.19	-0.85*** (0.31)	-1.46, -0.25
LE gain (QOL gain = reference)	-1.29*** (0.48)	-2.22, -0.35	-1.28 (1.22)	-3.66, 1.10	-1.22 (1.20)	-3.58, 1.14	-1.43 (0.90)	-3.20, 0.34
Age*LE gain (Age 10; QOL gain = reference)			-	-	-	-	-	-
20*LE gain	-	-	-1.13 (1.32)	-1.45, 3.71	1.08 (1.30)	-1.47, 3.63	1.86* (1.06)	-0.22, 3.93
40*LE gain			-0.25 (1.40)	-2.99, 2.49	-0.27 (1.38)	-2.98, 2.44	0.34 (1.11)	-1.83, 2.51
70*LE gain			-0.97 (1.28)	-3.49, 1.55	-0.86 (1.28)	-3.37, 1.66	-0.28 (0.99)	-2.22, 1.66
Severity[c] (10 = reference)	-	-	-	-	-	-	-	-
30	0.31 (0.46)	-0.59, 1.20	0.44 (0.46)	-0.47, 1.34	0.34 (0.45)	-0.54, 1.23	0.16 (0.39)	-0.60, 0.92
50	0.80** (0.42)	-0.02, 1.63	0.88** (0.42)	0.07, 1.69	0.90** (0.41)	0.08, 1.71	0.45 (0.35)	-0.22, 1.13
70	0.93** (0.40)	0.14, 1.72	0.92** (0.41)	0.12, 1.72	0.92** (0.40)	0.14, 1.69	0.39 (0.34)	-0.27, 1.04
90	1.73*** (0.45)	0.85, 2.61	1.77*** (0.46)	0.87, 2.66	1.76*** (0.45)	0.88, 2.64	1.07** (0.38)	0.32, 1.81
Health gain 50 points + death (Health gain 10 points + full recovery = reference)	2.20*** (0.45)	1.31, 3.09	2.13*** (0.46)	1.23, 3.03	2.08*** (0.45)	1.19, 2.97	1.62*** (0.39)	0.86, 2.38
Age (of respondents)	-	-	-0.29*** (0.11)	-0.51, -0.08	-0.28** (0.11)	-0.50, -0.06	-0.22** (0.10)	-0.42, -0.02
Age^2 (of respondents)	-	-	0.00** (0.00)	0.00, 0.01	0.00** (0.00)	0.00, 0.00	0.00* (0.00)	-0.00, 0.00
|Age respondents – Age patients|	-0.00 (0.01)	-0.02, 0.01	-	-	-	-	-	-
Sex (Female = reference)	0.78* 0.60)	-0.38, 1.95	0.80 (0.59)	-0.37, 1.96	0.80 (0.59)	-0.36, 1.96	1.22** (0.54)	0.17, 2.27
Children (No = reference)	1.31** (0.58)	0.17, 2.45	1.78*** (0.64)	0.52, 3.04	1.65** (0.64)	0.39, 2.91	1.42** (0.59)	0.26, 2.58
Education level[d] (Low = reference)	-	-	-	-	-	-	-	-
Medium	-0.13 (1.11)	-2.30, 2.04	-0.61 (1.16)	-2.89, 1.67	-0.28* 1.12)	-2.48, 1.92	-0.18 (1.00)	-2.13, 1.78
High	-0.38 (1.30)	-2.92, 2.16	-1.48 (1.34)	-4.10, 1.14	-0.77 (1.30)	-3.32, 1.79	-0.93 (1.13)	-3.14, 1.29
Household income[e] (Thousands)	1.61*** (0.42)	0.80, 2.42	1.57*** (0.42)	0.75, 2.40	1.78*** (0.42)	0.95, 2.62)	1.57*** (0.38)	0.82, 2.32
QOL (0-100 VAS)	-0.00 (0.02)	-0.04, 0.03	-0.02 (0.02)	-0.05, 0.02	-0.01 (0.02)	-0.04, 0.03)	0.01 (0.02)	-0.02, 0.04
CV task	-0.83*** (0.19)	-1.20, -0.45	-0.78*** (0.20)	-1.17, -0.39	-0.79*** (0.20)	-1.17, -0.41	-0.64*** (0.17)	-0.98, -0.31
WTP practice task[f]	-	-	0.01*** (0.00)	0.01, 0.02	-	-	-	-

	Model 5[A] Assess influence proximity of respondents' age to that of patients		Model 5[B] Corrected for WTP stated in practice task		Model 5[C] Excluding respondents with low clarity score (1−3) for practice task		Model 5[D] Minimum time to complete 39 seconds	
DV: WTP	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
Constant	4.46*** (1.42)	1.68, 7.23	12.99*** (2.94)	7.21, 18.77	11.52*** (2.94)	5.76, 17.29	11.05*** (2.71)	5.75, 16.35
R^2 overall	0.06		0.10		0.08		0.06	
n observations; groups	1,360; 468		1,332; 458		1,360; 468		1,783; 611	

CV, contingent valuation; DV, dependent variable; LE, life expectancy; NS, Not Stated; QOL, health-related quality of life (measured in points on a VAS); VAS, visual analogue scale (ranging from 0 "dead" to 100 "full health"); WTP, willingness to pay (in € per month for the duration of one year); [a] Note that sensitivity analyses were performed on all models (i.e. models 1−5). We present the results of the analyses that we performed based on model 5 as this model is most comprehensive and the analyses performed on the other models showed consistent results; [b] Age at onset of the disease (in years); [c] Severity is operationalised in terms of disease-related QOL loss and measured in points from 100 on the VAS. Note that the Severity 50 coefficients represent the impact on WTP of disease severity level 50 of patients who fully recovered or who died one year after falling ill; [d] Low = lower vocational and primary school, medium = middle vocational and secondary school, high = higher vocational and academic education; [e] Household income is adjusted for household size and calculated as household income / (household size)^0.5 to account for economies of scale [12]; [f] Respondents who gave protest zero valuations in the practice task are excluded from this table. * p<0.10, ** p<0.05, *** p<0.01.

Chapter 6
Who should receive treatment? An empirical enquiry into the relationship between societal views and preferences concerning healthcare priority setting

Based on: 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.

Abstract

Policymakers increasingly need to prioritise between competing health technologies or patient populations. When aiming to align allocation decisions with societal preferences, knowledge and operationalisation of such preferences is indispensable. This study examines the distribution of three views on healthcare priority setting in the Netherlands, labelled "Equal right to healthcare", "Limits to healthcare", and "Effective and efficient healthcare", and their relationship with preferences in willingness to trade-off (WTT) exercises. A survey including four reimbursement scenarios was conducted in a representative sample of the adult population in the Netherlands (n=261). Respondents were matched to one of the three views based on agreement with 14 statements on principles for resource allocation. We tested for WTT differences between respondents with different views and applied logit regression models for examining the relationship between preferences and background characteristics, including views. Nearly 65% of respondents held the view "Equal right to healthcare", followed by "Limits to healthcare" (22.5%), and "Effective and efficient healthcare" (7.1%). Most respondents (75.9%) expressed WTT in at least one scenario and preferred gains in quality of life over life expectancy, maximising gains over limiting inequality, treating children over elderly, and those with adversity over those with an unhealthy lifestyle. Various background characteristics, including the views, were associated with respondents' preferences. Most respondents held an egalitarian view on priority setting, yet the majority was willing to prioritise regardless of their view. Societal views and preferences concerning healthcare priority setting are related. However, respondents' views influence preferences differently in different reimbursement scenarios. As societal views and preferences are heterogeneous and may conflict, aligning allocation decisions with societal preferences remains challenging and any decision may be expected to receive opposition from some group in society.

6.1 Introduction

Healthcare resources are scarce and policymakers in publicly funded healthcare systems are increasingly confronted with the need to prioritise between competing health technologies or patient populations for reimbursement [1,2]. An important objective of a healthcare system is to generate as much health as possible, given the budget constraint [3]. To achieve this objective, economic evaluations of (new) health technologies are applied to guide policymakers in making decisions concerning the allocation of healthcare resources [4,5]. In health-economic evaluations, the value of a health technology is commonly expressed in terms of quality-adjusted life-years (QALYs) and evaluated against some monetary threshold value per QALY gained [1,6].

Regardless of whether the decision rule for economic evaluations implies the maximisation of health under a fixed budget or the maximisation of welfare for society, traditionally, health technologies with lower incremental cost-effectiveness ratios (ICERs) than some relevant threshold are eligible for funding [3,4,7]. Often, both the weight attached to QALY gains and the applied thresholds are constant in such evaluations. This assumes that a "QALY is a QALY is a QALY", regardless of beneficiary characteristics and the context in which QALYs are gained [8]. However, the practice of valuing all QALY gains equally, and hence regardless of these aspects, has become a matter of debate as evidence is accumulating that this may insufficiently reflect societal preferences [1,4,6,9–12]. Indeed, the public also considers an equitable or fair allocation of health and healthcare important in the allocation of healthcare resources [4,10,13,14] and societal preferences concerning healthcare priority setting are related to the (i) characteristics of healthcare beneficiaries, e.g. a patient's age, potential to benefit from treatment, remaining life-years, social role, and lifestyle, (ii) characteristics of the disease, e.g. the rarity of a disease and the burden of illness associated with a disease prior to treatment, and (iii) characteristics of interventions, e.g. the size, type, duration, and costs of health gains [1,4,10,15–18]. Although health economists tend to agree that such preferences should play a role in decisions concerning resource allocation in healthcare [4], they are generally not included in health-economic evaluations (even though notable exceptions like in the Netherlands exist [19]). The discrepancy between prioritisation based on health-economic evaluations and societal preferences for distributing health and healthcare is considered one of the reasons for the modest impact of health-economic evaluations on the outcome of allocation decisions [20–22]. To bridge this gap, knowledge and operationalisation of an equity-dependent decision rule appears to be indispensable.

Empirical evidence suggests that, although some members of the public appear unwilling to prioritise in healthcare, the majority accepts priority setting as being necessary [4,18,23]. However, little is known about the criteria that should be used according to the public and about the weight these should

Text box 6.1 Societal viewpoints on healthcare priority setting in the Netherlands

The view "Equal right to healthcare" comprises an egalitarian view on health and healthcare. People with this view consider access to healthcare a basic human right. Everyone is equal, hence has an equal right to healthcare. According to people with this view, prioritisation should solely be based on the need for care and prioritisation based on patient, disease, and intervention characteristics, such as the effect of treatment, is opposed. What is considered to be "the right care" is a matter of personal concern for patients and, according to people with this view, patients should be supported in their treatment choices regardless of the costs.

The view "Limits to healthcare" comprises a view with a strong concern for providing "the right care" for patients. People with this view consider health-related quality of life to be an important outcome of treatment. According to people with this view, providing the right care may imply refraining from (life prolonging) treatment. People with this view do not consider cost-effectiveness to be an important criterion for priority setting, although they do consider it important to make good use of money. Hence, providing treatments that generate minimal benefits should be avoided. Priority setting based on patient characteristics is rejected, with an exception made for lifestyle. According to people with this view, patients who are culpable of their own disease should receive lower priority and prevention should receive higher priority in allocation decisions.

The view "Effective and efficient healthcare" comprises a utilitarian view on health and healthcare. People with this view consider it important to generate as much health for society as possible given the budget constraint, and consider a patient's capacity to benefit from treatment important when setting priorities. Although people with this view focus on the cost-effectiveness of treatments, they do believe it is not possible to "put a [fixed] price on life". The value of health benefits depends on circumstances and patient characteristics, such as age and culpability, and hence these should be taken into account in priority setting.

receive in allocation decisions [17,18,24–27]. Commonly, studies examine societal preferences for priority setting on an aggregate or mean level [28]. Less common are studies that examine the heterogeneity of societal preferences or the relationship between underlying rationales and preferences [17,18,25–27]. In a previous study, Wouters et al. [17] used Q methodology to identify three societal viewpoints regarding healthcare priority setting among members of the public in the Netherlands: "Equal right to healthcare", "Limits to healthcare", and "Effective and efficient healthcare". Brief descriptions of these viewpoints can be found in Text box 6.1. In the current study, we examine the distribution of the three views in the general adult population and the relationship between these views and preferences concerning healthcare priority setting in four willingness to trade-off (WTT) exercises to inform priority-setting decisions in healthcare.

6.2 Methods

6.2.1 Sample and data collection

A professional internet survey company in the Netherlands distributed the questionnaire in October and November 2015, to a random sample that was stratified in terms of age, sex, and education level in order for it to be representative of the general adult population in the Netherlands regarding those characteristics. According to the Medical Research Involving Human Subjects Act, no ethical approval was required for this study. Prior to participating in the study, respondents were informed about the objectives of the study and how anonymity of respondents was guaranteed. They were informed that participation in the study was voluntary and could be stopped at any time, in which case the data they had provided would be discarded. Respondents could only enter the study after giving written consent for the use of their data for the purpose of the study.

Before answering questions about distributive preferences, respondents were explained that healthcare resources are scarce and that health policymakers inevitably have to make difficult choices between competing health technologies or patient populations for reimbursement. It was explained that the consequence of reimbursing one (type of) technology for one patient group implied not being able to reimburse another. Subsequently, respondents were asked to advise health policymakers on what would be the optimal allocation of available healthcare resources in four reimbursement scenarios.

In the next subsection, the questionnaire, including the statements and reimbursement scenarios, that was used for matchings respondents to a view and for eliciting their preferences is described. Subsequently, the reimbursement scenarios, scenario characteristics, and accompanying WTT exercises are discussed in more detail. In the final subsection of the Methods section, the analyses and hypotheses are described that were used for examining the distribution of the views and the relationship between respondents' views and preferences concerning healthcare priority setting.

6.2.2 Questionnaire

The questionnaire consisted of three parts. In part one, respondents were asked about demographic and background characteristics. In part two, respondents were asked to express their level of agreement on a seven-point Likert scale (ranging from completely disagree to completely agree) with 14 statements on principles for resource allocation that were extracted from Wouters et al. [17] and presented to respondents in random order. Table 6.1 presents these statements including respondents' mean (SD) level of agreement with each of the statements. To match respondents to one of the three views on healthcare priority setting, four statements were selected for each of the views and two additional statements were selected to untie, in case

Table 6.1 Overview of statements used for matching respondents to one of three societal views on healthcare priority setting (n=261)

			Factor score[a]			
Viewpoint	#	Statement	F1	F2	F3	Mean (SD)[b]
Equal right to healthcare	1	If it is possible to save a life, every effort should be made to do so	+3*	0	-2	5.21 (1.57)
	2	If there is a way of helping patients, it is morally wrong to deny them this treatment	+3*	+1	+1	5.38 (1.42)
	3	It's important to respect the wishes of patients who feel they should take every opportunity to extend their life	+1*	-3	-1	4.90 (1.35)
	4	Patient characteristics other than their health should play no role in prioritising care	+3*	0	-1	5.18 (1.46)
Limits to healthcare	5	At the end of life it is more important to provide a death with dignity than treatments that will only extend life for a short period of time	+2	+4*	+2	5.17 (1.50)
	6	People should accept that if it's your time to die, it's your time to die	0	+3*	0	4.48 (1.64)
	7	People who are in some way responsible for their own illness should receive lower priority than people who are ill through no fault of their own	-2*	+2*	0*	3.49 (1.70)
	8	There is no sense in saving lives if the quality of those lives will be really bad	0	+4*	-2	4.24 (1.63)
Effective and efficient healthcare	9	Children's health should be given priority over adult's health	-1*	-3*	+4*	3.78 (1.58)
	10	Priority should be given to patients who benefit most from treatment	-1*	+1*	+4*	4.07 (1.58)
	11	Priority should be given to those treatments that generate the most health	+1	+1	+3*	4.35 (1.57)
	12	Treatments that are very costly in relation to their health benefits should be withheld	-2*	0*	+1*	3.03 (1.56)
Additional statements	13	Treating people at the end of life is important, even if it is not going to result in big health gains	+1*	-1	-3	4.85 (1.50)
	14	Treating terminally ill patients as more 'worthy' of receiving care undervalues the health of other patients	0*	-1*	+1*	3.33 (1.67)

[a] Factor scores and p-values are extracted from Wouters et al. [17]; factor scores range from -4 (disagree most) to +4 (agree most); * p-value <0.01; F1 relates to the view "Equal right to healthcare", F2 to the view "Limits to healthcare", and F3 to the view "Effective and efficient healthcare"; b Respondents' mean (SD) level of agreement with the statements, expressed on a seven-point Likert scale ranging from 1 (completely disagree) to 7 (completely agree).

a respondent scored similarly on more than one view. The statements were selected based on the criteria that a statement should be characterising and distinguishing for one of the three views, which means that the statement should have a high factor score in that view and/or that this score should be statistically significantly different from factor scores of the other two views [17,29,30]. The assumption underlying the matching of respondents to one of the views was that respondents who expressed a relatively high level of agreement with statements that are characteristic and/or distinguishing for a specific view have a view that is similar to that view.

In part three of the questionnaire, respondents were presented four reimbursement scenarios. The scenarios were based on the study by Wouters et al. [17] and designed in such a way that differences in preferences between respondents with different views could manifest themselves. Each of the scenarios included two options, labelled A and B, that differentiated two competing treatments, based on the type of health gain, or patient groups, based on patients' potential to benefit from treatment, age, or lifestyle. Respondents were asked to advise health policymakers on reimbursement, by first choosing between the two treatments or patient groups and subsequently, depending on the scenario, indicating the relative size of the health gain or patient group that would make them indifferent between the two options. Respondents were allowed to opt out in case they had no preference for one of the options. When a respondent chose to opt out, they were asked to explain their choice by checking one of two provided answer options that indicated equality between the treatments or patient groups, e.g. 'both treatments are equally effective' or 'both treatments are equally ineffective', or by completing an open text field. As an example, scenario one is included in the supporting information.

6.2.3 Reimbursement scenarios

The reimbursement scenarios were similarly structured but differed in terms of treatment and patient characteristics. In scenario 1, respondents were asked to choose between two treatments based on their preference for health gains in terms of 3 points in health-related quality of life (QOL) or 3 months in life expectancy (LE), while both patient groups currently had a remaining LE of 3 months with a QOL of 3 points. The QOL scale ranged from 0 to 10, with 0 representing 'the worst imaginable health state' and 10 representing 'the best imaginable health state'. When a respondent preferred treatment A with a gain in QOL, they were asked at which point they would be indifferent between a gain in QOL between 0 and 3 points and a gain of 3 months in LE for treatment B (and vice versa if respondents preferred the LE gain). In scenario 2, respondents were asked to choose between two patient groups based on patients' potential to benefit from treatment. Respondents stated their preference for maximising health gains or limiting health inequality between the patient groups, by choosing between a 3 point gain in QOL in patient group A

or a 1 point gain in QOL in patient group B, while both groups currently had a QOL of 5 (on a scale from 0 to 10). When a respondent preferred the health maximising option, they were asked at which point they would be indifferent between a gain in QOL between 0 and 3 points and a gain of 1 point in QOL for the other patient group (or how large the difference should be, up to 5 points, to switch to patient group A, if they had a preference for patient group B). In scenario 3, respondents were asked to prioritise a 12 month increase in LE by choosing between two patient groups based on their preference for treating children (<18 years) or elderly (>70 years). When a respondent chose to treat the group of children, they were asked at how many months between 0 and 12 months gain in LE they would be indifferent between treatment of the two age groups (and vice versa if respondents preferred to treat the group of elderly). In scenario 4, respondents were asked to choose between two patient groups based on their preference for reducing the risk from 1:1,000 to 1:10,000 of a life-threatening illness for those with an unhealthy lifestyle or those with running the same risk due to adversity (explained to respondents as a reduction from 10 to 1 patients in a population of 10,000). When respondents preferred the patients running the risk due to adversity, they were asked to indicate how many patients between 1 and 10 would make them indifferent between the two groups (and vice versa). The scenarios stated there were no other differences between the treatments or patient groups than the ones described.

In all but scenario 2, the post-treatment health status was equal for patient groups in both options. However, in scenario 1, patients' post-treatment health status was not measured on a single scale, as in the other scenarios, but on a combined QOL and LE scale. In this scenario, the post-treatment health status was 18 for both patient groups. This was calculated in option A by multiplying 3 months LE by 6 points QOL and in option B by multiplying 6 months LE by 3 points QOL. Although in scenario 1 respondents' preferences were elicited on a combined QOL and LE scale, respondents who expressed a preference for a gain in QOL indicated their point of indifference on a QOL scale, while those who expressed a preference for a gain in LE indicated their point of indifference on a LE scale. Scenarios 2, 3, and 4 elicited respondents' point of indifference between the options on a single scale, either in terms of QOL, LE, or in number of patients.

6.2.4 Statistical analyses and hypotheses

To improve our sample's representativeness of the general adult population in the Netherlands, we weighted the data by applying a combined weighting factor for age, sex, and education level. For the analyses, respondents were divided into a 'traders' and a 'non-traders' subsample. Respondents who expressed WTT in at least one of the four scenarios were classified as 'trader' and those who did not express WTT were classified as 'non-trader'. Demographic and background characteristics of the sample and the two subsamples

were calculated in percentages of total and mean (SD). To match respondents to one of the views identified by Wouters et al. [17], we applied the following procedure. First, respondents' levels of agreement with the four statements were summed for each of the views and, for ease of interpretation, rescaled to a 0–10 scale. Next, respondents were matched to the view with the highest sum score on the condition that this score was above 5.0, hence indicated agreement. When two or three views received an equal highest sum score, the levels of agreement with statements 13 and 14 (see Table 6.1) were used to untie the scores and, if possible, used to match respondents to one of the views. Differences in characteristics between respondents who could and could not be matched, between traders and non-traders, and between respondents with different views were examined using independent t-, analysis of variance (ANOVA), and Fisher's exact tests. A Bonferroni correction was applied to adjust for the increased risk of a Type 1 error, caused by multiple comparisons.

In each scenario, the WTT of respondents between treatments or patient groups was examined by calculating the percentage of traders and non-traders, and the median (interquartile) range of indifference points of traders with a preference for option A or B. Differences in WTT and in median indifference points of traders with different views were examined using Fisher's exact and Kruskal-Wallis tests (Bonferroni corrected). Reasons of non-traders for opting out were explored qualitatively. To relate respondents' preferences to background characteristics, including view, logit regression models were applied. First, an overall model was composed for the four scenarios. This model included the variables age, age squared (to account for non-linearity), sex, education level, having children, daily smoking, and view. Having children and daily smoking were included as these variables were expected to be associated with the outcomes of interest in scenario 3 and 4. Subsequently, we applied likelihood ratio tests (LRT) to examine if this overall model could be improved for specific scenarios by including additional variables that might also be associated with the outcomes of interest, such as excessive alcohol use and being religious. We used generalised variation inflation factors (VIFs) to examine if the coefficient estimates' standard errors were inflated by multicollinearity.

Based on the description of the views in Wouters et al. [17], three hypotheses were formulated for the relationship between respondent's views and preferences concerning healthcare priority setting:

Hypothesis 1: Respondents with the view "Equal right to healthcare" have a lower WTT in all scenarios than respondents with the views "Limits to healthcare" and "Effective and efficient healthcare".

Hypothesis 2: The view "Limits to healthcare" is positively associated with respondents' WTT in all scenarios. In addition, respondents with this view

express a preference for health gains in terms of QOL in scenario 1, for health maximisation in scenario 2, and for treating those with adversity in scenario 4.

Hypothesis 3: The view "Effective and efficient healthcare" is positively associated with respondents' WTT in all scenarios. In addition, respondents with this view express a preference for health maximisation in scenario 2, for treating children in scenario 3, and for treating those with adversity in scenario 4.

The analyses were conducted using IBM SPSS Statistics 23.0 (SPSS, Inc, Chicago, Ill., USA) and Rstudio 0.99.903 (Rstudio, Inc., Boston, MA, USA).

6.3 Results

The data were weighted by applying a combined weighting factor for age, sex, and education level with a mean (SD) of 1.00 (0.47). Table 6.2 presents the descriptive statistics and the distribution of the views in the weighted sample (n=261), and in the traders and non-traders subsamples. The majority of respondents (n=198; 75.9%) expressed WTT in at least one scenario. Of the respondents, 90.2% (n=235) could be matched to one of the views based on their level of agreement with the 12 statements and 3.9% (n=10) could be matched based on their level of agreement with the two additional statements. A t-test revealed that respondents who could not be matched to one of the views were relatively younger than those who could be matched. Mean (SD) age of respondents who could not be matched was 32.3 (13.8) years and of those could be matched 47.2 (14.7) years. This difference was significant at the 0.01 level (two-tailed, Bonferroni corrected, $\alpha/12$). In addition, a Fisher's exact test revealed that the difference in nationality between respondents who could and could not be matched was also significant at the 0.01 level (two-tailed, Bonferroni corrected, $\alpha/12$). Respondents with a Dutch nationality could more frequently be matched to a view than respondents with a different nationality.

The majority of respondents was matched to the view "Equal right to healthcare" (64.5%), followed by "Limits to healthcare" (22.5%), and "Effective and efficient healthcare" (7.1%). A minority of respondents (5.9%) could not be matched. A similar distribution of views was observed among traders (60.1%, 28.8%, 7.6, and 3.5% respectively). However, among non-traders the view "Equal right to healthcare" was considerably more prevalent (78.3%), while the views "Limits to healthcare" and "Effective and efficient healthcare" were less prevalent (2.9% and 5.4%, respectively), and relatively more non-traders could not be matched (13.4%). A Fisher's exact test (two-tailed, Bonferroni corrected, $\alpha/13$) revealed that the difference in views between traders and non-traders was significant at the 0.001 level. In addition, a Fisher's exact test (two-tailed, Bonferroni corrected, $\alpha/13$) revealed that traders were more frequently highly educated and less frequently smoked daily (p-value <0.05) than non-traders. No differences were revealed between traders and non-tra-

Table 6.2 Sample characteristics (n=261)[a]

	Total (n=261)		Traders (n=198)		Non-traders (n=63)		p-value
	%	Mean (SD)	%	Mean (SD)	%	Mean (SD)	
Age (Years)		46.3 (15.1)		47.2 (15.1)		43.5 (14.9)	0.087
Sex (Female)	49.4		52.7		44.0		0.312
Nationality (Dutch)	88.9		88.4		90.5		0.008
Education level[b]							0.003*
Low	23.9		19.0		39.4		
Middle	50.7		52.6		44.7		
High	25.4		28.5		15.9		
Living situation							0.023
Single	27.0		24.6		34.6		
Married/cohabitant	63.0		67.3		49.2		
With others	9.5		7.3		16.2		
Children (Yes)	60.1		64.3		46.5		0.012
Lifestyle							
Smoking (Daily)	17.1		12.0		32.9		0.001*
Alcohol usage (Excessive)[c]	20.5		21.3		17.9		0.717
Chronic condition (Yes)							
Physical	31.7		31.4		32.8		0.551
Mental	4.6		5.0		3.4		
Physical and mental	2.7		3.5		0.0		
Religious (Yes)[d]	26.8		27.2		25.4		0.871
View on healthcare priority setting							0.000***
Equal right to healthcare	64.5		60.1		78.3		
Limits to healthcare	22.5		28.8		2.9		
Effective and efficient healthcare	7.1		7.6		5.4		
Not matched	5.9		3.5		13.4		
Health status (VAS 0-10)		6.8 (1.5)		6.8 (1.6)		7.0 (1.4)	0.337
Happiness (VAS 0-10)		7.2 (1.6)		7.2 (1.6)		7.2 (1.8)	0.905

VAS, Visual Analogue Scale; [a] In this table, respondents who expressed willingness to trade-off (WTT) in at least one reimbursement scenario are classified as "traders", respondents who expressed no WTT in all four reimbursement scenarios are classified as "non-traders"; [b] Low = lower vocational and primary school, Middle = middle vocational and secondary school, High = higher vocational and academic education; [c] Applied standard for excessive alcohol use for female respondents: consumption of ≥7 alcohol units per week or of ≥4 alcohol units on one day, for male respondents: consumption of ≥14 alcohol units per week or ≥6 alcohol units on one day; [d] Operationalised by the question "Do you consider yourself to be part of a religious community (yes/no)?"; * $p<0.05$, *** $p<0.001$ (two-tailed, Bonferroni corrected, $\alpha/13$).

Table 6.3 Willingness to trade-off (in n and % of total), scenario (S) and option (A and B) specifications, and traders' preferences (in % of total) and median (IQR) indifference point (n=261)

S	WTT n (%)	Option	Specification	Pre-treatment health status			Treatment benefit			Post-treatment health status[a]		Preference (%)	Median (IQR) indifference point
				QOL[b]	LE[c]	Risk	QOL[b]	LE[c]	Risk	QOL[b]	LE[c]		
1	132 (50.8)	A	QOL	3	3	NS	3	NS	NS	6	3	80.3	1.5 (1.0-2.0)
		B	LE	3	3	NS	NS	3	NS	3	6	19.7	2.0 (1.0-2.5)
2	109 (42.0)	A	Maximize health	5	NS	NS	3	NS	NS	8	NS	91.7	2.0 (1.5-2.0)
		B	Limit inequality	5	NS	NS	1	NS	NS	6	NS	8.3	4.0 (3.9-4.5)
3	123 (47.1)	A	Children	NS	NS	NS	NS	12	NS	NS	12	94.3	6.0 (4.0-10.0)
		B	Elderly	NS	NS	NS	NS	12	NS	NS	12	5.7	9.3 (6.0-10.0)
4	117 (44.8)	A	Lifestyle-related risk	NS	NS	1:1,000	NS	NS	1:10,000	NA	NA	10.3	7.0 (5.5-7.6)
		B	Adversity	NS	NS	1:1,000	NS	NS	1:10,000	NA	NA	89.7	5.0 (4.0-8.0)

IQR = interquartile range; LE, life expectancy; NA = Not Applicable; NS = Not Stated, i.e. equal for both options; QOL, health-related quality of life; WTT, willingness to trade-off; [a] Post-treatment health status is calculated by aggregating patients' pre-treatment health status and the treatment benefit; [b] QOL is noted in points on a 0−10 scale; [c] LE is noted in months.

ders concerning other characteristics. Between respondents with different views, a Fisher's exact test (two-tailed, Bonferroni corrected, α/12) revealed a significant difference at the 0.05 level for education level (not in table). Respondents with the view "Equal right to healthcare" were more frequently lower educated than respondents with the views "Limits to healthcare" and "Effective and efficient healthcare".

Table 6.3 presents the scenario specifications, the proportion of respondents who were willing to trade-off, respondents' preferences for option A or B, and their median (IQR) indifference points for each of the scenarios. Although the distribution of indifference points is different between traders with a preference for option A or B, an overlap of IQR can be seen in scenario's 3 and 4. Note that for scenario 1 the IQR for option A and B are on a different scale.

The percentage of respondents who were willing to trade-off ranged between 42.0% and 50.8% in the four scenarios. The highest WTT percentage was expressed in scenario 1, where respondents were asked to prioritise between health gains in terms of QOL or LE. The lowest WTT percentage was expressed in scenario 2, where respondents were asked to prioritise between maximising health gains and limiting health inequality between patient groups. Of the traders, a large majority expressed a preference for health gains in terms of QOL (80.3%) over gains in LE (19.7%), for maximising health gains (91.7%) over limiting health inequality (8.3%), for treating children (94.3%) over treating elderly (5.7%), and reducing risk for those with adversity (89.7%) over those with an unhealthy lifestyle (10.3%). In each of the scenarios, 81.0% to 84.0% of non-traders consisted of respondents with the view "Equal right to healthcare", who opted out more frequently than respondents with the views "Limits to healthcare" and "Effective and efficient healthcare". Table 6.4 presents the differences in WTT between respondents with different views in each of the scenario. Fisher's exact tests revealed that these differences were significant at the 0.001 level in scenario 1, 3, and 4, and at the 0.01 level in scenario 2 (two-tailed, Bonferroni corrected, α/4). The difference in the median indifference points of respondents with different views was not significant.

Between 79.8% and 92.9% of non-traders explained their preference for opting out by checking one of the provided answer options, the remainder by completing the open text field. In scenario 1, 20.2% (n=26) of the non-traders completed the open text field of which 69.2% stated that the choice between options A and B was not theirs but only for patients themselves to make. For example, because "having a preference for quality of life or life expectancy is a personal matter". Other explanations for opting out included "both options are very much alike" or "both patient groups will die regardless of treatment". In scenario 2, 10.9% (n=16) completed the open text field. Explanations for opting out included "I do not see a difference between the two options", "I would treat whoever came first", "the value of a person's life cannot solely

Table 6.4 Willingness to trade-off of respondents with different views on healthcare priority setting (n=246)[a]

View	WTT in scenario[b]				No WTT in any scenario
	1	2	3	4	
Equal right to healthcare	74	59	64	58	49
Limits to healthcare	45	36	42	43	2
Effective and efficient healthcare	10	11	11	13	3
n	129	106	117	114	62

WTT, willingness to trade-off (in frequencies); [a] Respondents who could not be matched to one of the views (n=15) are excluded from the table; [b] In scenario 1, respondents expressed WTT by choosing for a gain in quality of life or in life expectancy or expressed no WTT by opting out. In scenario 2, respondents expressed WTT by choosing for health maximization or limiting health inequality or expressed no WTT by opting out. In scenario 3, respondents expressed WTT by choosing for treating children or elderly or expressed no WTT by opting out. In scenario 4, respondents expressed WTT by choosing for treating those with an unhealthy lifestyle or adversity or expressed no WTT by opting out. Respondents who opted out in all four scenarios are included in the table under "no WTT in any scenario". The presented differences in WTT frequencies are significant at the 0.001 level in scenario 1, 3, and 4, and at the 0.01 level in scenario 2 (two-tailed, Bonferroni corrected, α/4).

be determined based on the physical condition of that person", and "quality of life is an abstract concept and it provides too little information to form an informed opinion". In scenario 3, 7.1% (n=10) completed the open text field. Explanations for opting out included "although I have a preference for treating children, the age of patients should not matter", "my preference in this matter depends entirely on the burden of illness of the patients", and "quality of life matters more than life expectancy". In scenario 4, 8.7% (n=12) completed the open text field and stated, for example, "it is nearly impossible to determine whether a person is culpable of their own disease", "two-third of all cancer cases are caused by having bad luck", "having an unhealthy lifestyle may be involuntary", and "having an unhealthy lifestyle is often due to adversity". The explanations for opting out did not seem to differ between respondents with different views on healthcare priority setting.

Table 6.5 and 6.6 present the results of the logit regression models examining the relationship between background characteristics and the WTT, and the most preferred option of respondents in each of the scenarios, with the baseline set to a preference for opting out.

The results of the LRTs indicated that the overall model could not be significantly improved for any of the scenarios by including additional variables. Hence, logit regression models with the same independent variables are presented for all four scenarios. The VIFs indicated no multicollinearity (VIFs <1.90) for all variables except for age and age squared (VIF 40.78–46.60). The higher VIFs for age and age squared can be explained by the correlation between these two variables. When excluding age or age squared from the regression models, the corresponding VIFs were all <1.62.

In terms of background characteristics, having a higher age (OR 0.845–0.877), having children (OR 0.261), and daily smoking (OR 0.219–0.312) negatively affected, and having a high education level (OR 2.865–3.072) positively affected the WTT of respondents in different reimbursement scenarios. Having a higher age (OR 0.872), having children (OR 0.251), and daily smoking (OR 0.296) were negatively associated, and having a high education level (OR 3.122) was positively associated with a preference for treating children. Having a higher age (OR 0.870) and daily smoking (OR 0.252) were also negatively associated with a preference for treating those with adversity. In addition, daily smoking (OR 0.261) was negatively associated with a preference for health maximisation. Having a middle or high education level (OR 2.745–5.309) was positively associated with a preference for health gains in terms of QOL. Compared to the view "Equal right to healthcare", the views "Limits to healthcare" (OR 3.306–5.850) and "Effective and efficient healthcare" (OR 2.608–4.375) were positively associated with the WTT of respondents. The view "Limits to healthcare" was also positively associated with a preference for health gains in terms of QOL (OR 4.241), maximising health gains (OR 3.443), treating children (OR 5.354), and those with adversity (OR 6.443). The view "Effective and efficient healthcare" was positively associated with a preference for health maximisation, treating children, and those with adversity (OR 2.640–4.050).

The WTT of respondents with different views on healthcare priority setting differed significantly in each of the scenarios and the majority of non-traders in each of the scenarios consisted of respondents with the view "Equal right to healthcare". These findings provide evidence in support of hypothesis 1. The logit regression analyses discussed above provide additional evidence in support of hypothesis 1 by indicating that, compared to having the view "Equal right to healthcare", the views "Limits to healthcare" and "Effective and efficient healthcare" are positively associated with WTT and the most preferred option in all scenarios. The logit regression analyses also provide evidence in support of hypotheses 2 and 3. Having the view "Limits to healthcare" was positively associated with WTT in all scenarios as well as with a preference for health gains in terms of QOL, health maximisation, and treating those with adversity. Having the view "Effective and efficient healthcare" was also positively associated with WTT in all scenarios. In addition, having this view was positively associated with a preference for health maximisation, treating children, and those with adversity. The lower significance levels that accompany these latter associations may be explained by the relatively small number of respondents having the view "Effective and efficient healthcare" (n=18).

Table 6.5 Impact of characteristics on the willingness to trade-off (WTT) yes/no of respondents in four reimbursement scenarios (logit regression model, n=246)[a]

	Scenario 1[b]		Scenario 2[c]		Scenario 3[d]		Scenario 4[e]	
	β (SE)	OR (95% CI)	β (SE)	OR (95% CI)	β (SE)	OR (95% CI)	β (SE)	OR (95% CI)
Age	0.050 (0.061)	1.051 (0.934,1.187)	-0.078 (0.059)	0.925 (0.823,1.039)	-0.168** (0.063)	0.845 (0.745,0.955)	-0.131* (0.062)	0.877 (0.776,0.988)
Age^2	-0.000 (0.001)	1.000 (0.998,1.001)	0.001 (0.001)	1.001 (0.999,1.002)	0.002* (0.001)	1.002 (1.000,1.003)	0.001· (0.001)	1.001 (1.000,1.003)
Female	0.042 (0.284)	1.042 (0.598,1.823)	0.197 (0.286)	1.218 (0.695,2.142)	0.439 (0.298)	1.551 (0.867,2.799)	-0.038 (0.297)	0.962 (0.537,1.723)
Education level (low = reference)	-	-	-	-	-	-	-	-
Middle	0.560 (0.355)	1.750 (0.879,3.555)	0.267 (0.370)	1.306 (0.635,2.732)	0.442 (0.384)	1.556 (0.738,3.351)	0.413 (0.384)	1.511 (0.717,3.245)
High	1.122** (0.435)	3.072 (1.323,7.325)	0.315 (0.434)	1.371 (0.586,3.232)	1.053* (0.457)	2.865 (1.184,7.155)	0.368 (0.452)	1.445 (0.596,3.531)
Children (no = reference)	0.015 (0.350)	1.015 (0.511,2.022)	-0.199 (0.353)	0.820 (0.406,1.632)	-1.345*** (0.397)	0.261 (0.116,0.555)	-0.377 (0.370)	0.686 (0.328,1.406)
Smoking (not daily = reference)	-0.586 (0.381)	0.557 (0.258,1.158)	-1.164** (0.444)	0.312 (0.122,0.713)	-1.216** (0.453)	0.297 (0.115,0.691)	-1.520** (0.471)	0.219 (0.080,0.522)
View (Equal right to healthcare = reference)	-	-	-	-	-	-	-	-
Limits to healthcare	1.196** (0.376)	3.306 (1.614,7.102)	1.205*** (0.356)	3.336 (1.677,6.793)	1.641*** (0.398)	5.161 (2.421,11.606)	1.766*** (0.386)	5.850 (2.807,12.843)
Effective and efficient healthcare	0.337 (0.546)	1.401 (0.475,4.158)	0.959· (0.546)	2.608 (0.903,7.901)	0.931 (0.578)	2.536 (0.821,8.150)	1.476* (0.587)	4.375 (1.448,15.043)
Constant	-2.353 (1.622)	0.095 (0.004,2.179)	1.447 (1.595)	4.252 (0.188,100.414)	4.652** (1.717)	104.790 (3.820,3294.049)	2.830· (1.657)	16.950 (0.679,460.767)
AIC	338.81		338.43		311.72		324.33	
R^2 (Mc Fadden)	0.172		0.160		0.242		0.205	
Adjusted R^2 (McFadden)	0.121		0.107		0.190		0.152	

[a] Respondents who could not be matched to a view (n=15) are excluded from this analysis; [b] Scenario 1: WTT between health benefits in terms of quality of life or life expectancy (yes) or opt out (baseline); [c] Scenario 2: WTT between health maximisation or limiting health inequality between patient groups (yes) or opt out (baseline); [d] Scenario 3: WTT between children <18 years or elderly >70 years (yes) or opt out (baseline); [e] Scenario 4: WTT between patients with a risk of becoming ill due to having an unhealthy lifestyle or due to adversity (yes) or opt out (baseline); · p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 6.6 Impact of background characteristics of traders on preferences for choice of most preferred-option (A or B) in four reimbursement scenarios (logit regression model)[a]

	Scenario 1A[b]		Scenario 2A[c]		Scenario 3A[d]		Scenario 4B[e]	
	β (SE)	OR (95% CI)	β (SE)	OR (95% CI)	β (SE)	OR (95% CI)	β (SE)	OR (95% CI)
Age	0.106 (0.072)	1.111 (0.968,1.287)	-0.088 (0.061)	0.915 (0.811,1.032)	-0.137* (0.066)	0.872 (0.765,0.991)	-0.140* (0.063)	0.870 (0.768,0.982)
Age squared	-0.001 (0.001)	0.999 (0.998,1.0001)	0.001 (0.001)	1.001 (0.999,1.002)	0.001· (0.001)	1.001 (1.000,1.003)	0.001* (0.001)	1.001 (1.000,1.003)
Female	-0.067 (0.313)	0.936 (0.506,1.733)	0.247 (0.296)	1..281 (0.718,2.296)	0.397 (0.304)	1.488 (0.822,2.714)	0.193 (0.310)	1.213 (0.661,2.237)
Education level (low = reference)	-	-	-	-	-	-	-	-
Middle	1.010* (0.413)	2.745 (1.245,6.351)	0.151 (0.382)	1.163 (0.552,2.486)	0.469 (0.399)	1.599 (0.738,3.553)	0.377 (0.392)	1.458 (0.681,3.192)
High	1.669*** (0.491)	5.309 (2.070,14.322)	0.160 (0.450)	1.174 (0.484,2.853)	1.139* (0.469)	3.122 (1.261,8.000)	0.279 (0.465)	1.322 (0.531,3.312)
Children (no = reference)	0.094 (0.381)	1.098 (0.521,2.331)	-0.230 (0.368)	0.795 (0.382,1.625)	-1.384*** (0.402)	0.251 (0.111,0.539)	-0.224 (0.388)	0.800 (0.370,1.702)
Smoking (not daily = reference)	-0.439 (0.427)	0.644 (0.271,1.464)	-1.343** (0.491)	0.261 (0.090,0.641)	-1.218** (0.467)	0.296 (0.111,0.705)	-1.377** (0.475)	0.252 (0.092,0.607)
View (Equal right to healthcare = reference)	-	-	-	-	-	-	-	-
Limits to healthcare	1.445*** (0.390)	4.241 (2.012,9.370)	1.236*** (0.365)	3.443 (1.700,7.134)	1.678*** (0.340)	5.354 (2.503,12.088)	1.863*** (0.393)	6.443 (3.050,14.341)
Effective and efficient healthcare	0.514 (0.610)	1.671 (0.494,5.598)	0.971· (0.563)	2.640 (0.879,8.236)	1.049· (0.580)	2.854 (0.921,9.204)	1.399* (0.604)	4.050 (1.283,14.256)
Constant	-4.661* (1.911)	0.009 (0.000,0.357)	1.729 (1.646)	5.633 (0.227,148.336)	3.940* (1.781)	51.411 (1.630,1816,816)	2.401 (1.712)	11.033 (0.394,333.256)
AIC	284.72		318.58		297.68		306.94	
R^2 (Mc Fadden)	0.229		0.174		0.252		0.205	
Adjusted R^2 (McFadden)	0.171		0.118		0.198		0.150	

[a] Respondents who could not be matched to a view (n=15) are excluded from this analysis; [b] Scenario 1A = preference of traders for health benefit in terms of health-related quality of life (n=105), baseline = respondents who opted out (n=116); [c] Scenario 2A = preference of traders for health maximisation (n=97), baseline = respondents who opted out (n=138); [d] Scenario 3A = preference of traders for treating children ≤18 years (n=111), baseline = respondents who opted out (n=130); [e] Scenario 4B = preference of traders for treating persons with a risk of becoming ill due to adversity (n=102), baseline = respondents who opted out (n=131); · p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

6.4 Discussion

This study was performed against the background of the ongoing debate about societal concerns for an equitable and fair allocation of healthcare resources. The aim of this study was twofold. The first aim was to examine the distribution of three societal views on healthcare priority setting, i.e. "Equal right to healthcare", "Limits to healthcare", and "Effective and efficient healthcare" [17], in the general adult population in the Netherlands. The second aim was to examine the relationship between the views and preferences concerning healthcare priority setting, by examining respondents' WTT between treatments or patient groups in four different reimbursement scenarios as well as by relating respondents' preferences to background characteristics, including their view.

The results of our study suggest that "Equal right to healthcare" is the most prevalent view on healthcare priority setting in Dutch society. Based on our analyses, we found evidence in support of the hypothesis that respondents with this view had a lower WTT in the different reimbursement scenarios than respondents with the views "Limits to healthcare" and "Effective and efficient healthcare". In addition, we found evidence in support of hypotheses 2 and 3. The view "Limits to healthcare" is positively associated with WTT and with a preference for health gains in terms of QOL, health maximisation, and treating those with adversity. The view "Effective and efficient healthcare" is positively associated with WTT and with a preference for health maximisation, treating children, and reducing the risk of a life threatening disease for people with adversity. It should be noted, however, that the significance levels of these associations were higher for having the view "Limits to healthcare" than for having the view "Effective and efficient healthcare". Although, on average, the WTT differed between respondents with different views, our results suggest that the indifference points of those who are willing to trade-off did not differ, hence did not depend on their view. Our results also suggest that those who are willing to trade-off in different reimbursement scenarios generally prefer gains in QOL over LE, maximising health gains over limiting health inequality, treating children over the elderly, and treating those with adversity over those with an unhealthy lifestyle.

The finding that the majority of the general adult population in the Netherlands is willing to trade-off between competing health technologies or patient populations in at least one reimbursement scenario is in line with other empirical studies that suggest that the majority of the public is willing to prioritise in healthcare [4,23,24]. Although a large majority was willing to trade-off in at least one scenario, the proportions were close to 50% in each of the scenarios separately. Hence, respondents' characteristics and their view on healthcare priority setting influenced their WTT and preferences differently in different reimbursement scenarios. As preferences of the public are heterogeneous and may conflict, aligning reimbursement decisions with these preferences

is challenging and any allocation decision made by health policymakers may receive opposition from some group in society.

Empirical evidence about the heterogeneity of societal preferences or the relationship between underlying rationales and preferences for priority setting is limited [17,18,25–27], and research relating views on healthcare priority setting to such preferences may indeed be considered innovative. Using the same methodology as Wouters et al. [17], Baker et al. [25,30] identified three views on healthcare priority setting in the United Kingdom (UK) and examined the distribution of these views in British society. McHugh et al. [26] identified three views on the relative value of end-of-life treatments and, in a more recent study; Mason et al. [27] examined the distribution of these views in British society. Although in both studies, two of the identified views share similarities with the views "Equal right to healthcare" and "Effective and efficient healthcare" of Wouters et al. [17], none of the views appeared to be as dominant in the UK [25,27] as the view "Equal right to healthcare" in the Netherlands. Van Exel et al. [18] identified five views on healthcare priority setting in ten European countries, among which the Netherlands and the UK. Mason et al. [31] examined the distribution of these views in a subset of nine countries. The results of this study support our finding that an egalitarian view on healthcare priority setting is the most prevalent view in the Netherlands. In addition, the results of this study suggest that an egalitarian view on healthcare priority setting in the most prevalent view in the UK as well as in the other European countries. Further comparative research will be necessary to investigate the difference in views and their distribution between countries, and the relationship between the views and societal preferences for priority setting in these countries. The results of our study generally align with the results of other studies indicating that the social value of the QALY does not exist [14] as societal concerns for an efficient and equitable allocation of health and healthcare are heterogeneous. Our findings support those of other studies indicating, for example, that priority should be given to younger over older people [32–36] and to those with adversity over those with an unhealthy lifestyle [37–39].

Some limitations of this study need to be mentioned. A first limitation concerns the four relatively simple WTT exercises. Because our primary aim was to explore the relationship between the three societal views and preferences in a number of reimbursement scenarios based on distinguishing characteristics of those views, we chose for fairly straightforward WTT exercises. In addition, we expected that respondents might find the WTT exercises rather difficult and, therefore, kept the WTT exercises clear and concise. A second limitation concerns the initial lack of representativeness of our sample, resulting from suboptimal recruitment of respondents. To improve our sample's representativeness, we weighted the data by applying a combined weighting factor for age, sex, and education level. Although this method is often associated with an increased level of uncertainty concerning the results, a comparison of the

results pre and post weighting indicated no major changes in the size or direction of estimates. A third limitation concerns the non-randomised order in which the reimbursement scenarios were presented to respondents. However, as respondents were presented with only four scenarios that differed in terms of treatment and patient characteristics, we expect the possible risk of order bias to be limited. A fourth limitation is concerned with the possibility for respondents to avoid prioritisation and opt out in each of the four scenarios. Although we provided an opt-out to examine respondents' WTT and to explore non-traders' reasons for opting out in each of the scenarios, in decision-making practice opting out is not possible for health policymakers. It is unclear, how not providing an opt-out would have influenced the results of our study. A fifth limitation concerns the exclusion of respondents that could not be matched from the logit regression analyses. These respondents significantly differed in age and nationality from respondents who could be matched and, excluding these respondents resulted in a loss of information concerning our sample's preferences in the different reimbursement scenarios. However, excluding these respondents did not affect our primary aim of conducting the regression analyses, i.e. to explore the relationship between the three views and respondents' preferences in different reimbursement scenarios. In addition, as the excluded group of respondents was relatively small (n=15) and could not be matched to one of the views, we considered the loss of information to be limited. A final limitation concerns the lack of a normative discussion of the views and preferences concerning healthcare priority setting. Our aim was to examine the distribution of the three views and the relationship between the views and preferences and, therefore, a normative discussion was outside the scope of this paper. We refer the interested reader to e.g. Schwappach [4], Olsen et al. [15], Ottersen [40], and Bognar [41,42] for normative discussions about societal preferences concerning healthcare priority setting. In addition to these limitations, we would like to address that we consider it a strength of our study that we combined different methods to examine the relationship between societal views on healthcare priority setting and preferences in different healthcare reimbursement scenarios. This type of study is regarded as methodologically challenging [30] and is infrequently conducted.

Our results indicate that societal preferences concerning healthcare priority setting are heterogeneous and complex as people's view on healthcare priority setting and background characteristics influence their preferences differently in different reimbursement scenarios. Hence, when aiming to align allocation decisions with societal preferences for equity and efficiency, the use of a mix of equity concerns in decision-making practice is recommended. As we examined the relationship between societal views and preferences in the context of only four reimbursement scenarios, we recommend extending our research to scenarios that include other potential sources that contribute to the social value of the QALY [4,14]. For example, patients' prior healthcare consumption, the duration of health benefits, and the burden of illness that is associated with a disease. As heterogeneous preferences may sometimes be conflic-

ting, aligning allocation decisions with societal preferences is challenging and decisions will almost inevitably receive opposition from some group or another in society. Given the available evidence, this is unlikely to be a strictly Dutch phenomenon. Hence, knowledge about the (distribution of the) societal views and related preferences concerning healthcare priority setting may help health policymakers to be considerate of these views and preferences when allocating resources in healthcare. This knowledge, for example about the high prevalence of the egalitarian view on healthcare priority setting, may also help health policymakers in communicating and explaining (inevitable) allocation decisions to the public.

6.5 Conclusions

The results of this study suggest that "Equal right to healthcare" is the most prevalent view on healthcare priority setting in the Netherlands. Although we expected this egalitarian view to be negatively associated with WTT, our results indicate that the majority of people is still willing to prioritise between competing health technologies or patient groups regardless of their view on healthcare priority setting. People's characteristics and views on healthcare priority setting influence preferences differently in different reimbursement scenarios. As societal views and preferences are heterogeneous and may conflict, aligning allocation decisions with societal preferences remains challenging and any decision may be expected to receive opposition from some group in society. When aiming to align allocation decisions with societal preferences concerning healthcare priority setting, accounting for the variety in societal views and preferences is recommended.

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 Health Econ. 2013;14(1):107–15.
2. Sabik LM, Lie RK. Priority setting in health care: Lessons from the experiences of eight countries. Int J Equity Health. 2008;7:1–13.
3. Gold M, Siegel J, Russell L, Weinstein M. Cost-Effectiveness in Health and Medicine. New York: Oxford University Press; 1996.
4. Schwappach DL. Resource allocation, social values and the QALY: A review of the debate and empirical evidence. Health Expect. 2002;5(3):210–22.
5. Schreyögg J, Stargardt T, Velasco-Garrido M, Busse R. Defining the "health benefit basket" in nine European countries: Evidence from the European Union Health BASKET Project. Eur J Health Econ. 2005;6:2–10.
6. 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):1119–26.
7. 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.
8. Weinstein MC. A QALY is a QALY is a QALY - Or is it? J Health Econ. 1988;7(3):289–90.
9. Nord E. The person-trade-off approach to valuing health care programs. Med Decis Mak. 1995;15(3):201–8.
10. 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.
11. Wagstaff A, Doorslaer E V. Measuring and Testing for Inequity in the Delivery of Health Care. J Hum Resour. 2000;35(4):716–33.
12. Cookson R, Drummond M, Weatherly H. Explicit incorporation of equity considerations into economic evaluation of public health interventions. Health Econ Policy Law. 2009;4(2):231–45.
13. Wagstaff A. QALYs and the equity-efficiency trade-off. J Health Econ. 1991;10(1):21–41.
14. Brouwer WBF, van Exel NJA, Baker R, Donaldson C. The New Myth: The Social Value of the QALY. Pharmacoeconomics. 2008;26(1):1–4.
15. 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):1163–72.
16. Shah KK. Severity of illness and priority setting in healthcare: A review of the literature. Health Policy. 2009;93(2–3):77–84.
17. Wouters S, van Exel J, Baker R, B.F. Brouwer W. Priority to End of Life Treatments? Views of the Public in the Netherlands. Value Health. 2017;20(1):107–17.
18. 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;126:128–37.

19. 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):621–9.
20. Rutten F, Van Busschbach J. How to define a basic package of health services for a tax funded or social insurance based health care system? HEPAC Health Econ Prev Care. 2001;2(2):45–6.
21. Stolk EA, Pickee SJ, Ament AHJA, Busschbach JJV. Equity in health care prioritisation: An empirical inquiry into social value. Health Policy. 2005;74(3):343–55.
22. Franken M, Nilsson F, Sandmann F, De Boer A, Koopmanschap M. Unravelling drug reimbursement outcomes: A comparative study of the role of pharmacoeconomic evidence in dutch and swedish reimbursement decision making. Pharmacoeconomics. 2013;31(9):781–97.
23. Menzel P, Gold MR, Nord E, Richardson J, Ubel P. Toward a Broader View of Values in Cost-Effectiveness Analysis of Health. :7–15.
24. Dolan P, Cookson R. A qualitative study of the extent to which health gain matters when choosing between groups of patients. Health Policy. 2000;51(1):19–30.
25. Baker R, Wildman J, Mason H, Donaldson C. Q-Ing For Health—A New Approach To Eliciting The Public's Views On Health Care Resource Allocation. Health Econ. 2014;23:283–97.
26. McHugh N, Baker RM, Mason H, Williamson L, Van Exel J, Deogaonkar R, et al. Extending life for people with a terminal illness: A moral right and an expensive death? Exploring societal perspectives. BMC Med Ethics. 2015;16(1):1–15.
27. Mason H, Collins M, McHugh N, Godwin J, Van Exel J, Donaldson C, et al. Is "end of life" a special case? Connecting Q with survey methods to measure societal support for views on the value of life-extending treatments. Health Econ. 2018;27(5):819–31.
28. van de Wetering L, van Exel J, Bobinac A, Brouwer WBF. Valuing QALYs in Relation to Equity Considerations Using a Discrete Choice Experiment. Pharmacoeconomics. 2015;33(12):1289–300.
29. van Exel J, de Graaf G, Brouwer W. Give me a break!. Informal caregiver attitudes towards respite care. Health Policy. 2008;88(1):73–87.
30. 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).
31. Mason H, van Exel J, Baker R, Brouwer W, Donaldson C, Pennington M, et al. From representing views to representativeness of views: Illustrating a new (Q2S) approach in the context of health care priority setting in nine European countries. Soc Sci Med. 2016;166:205–13.
32. Lewis PA, Charny M. Which of two individuals do you treat when only their ages are different and you can't treat both? J Med Ethics. 1989;15(1):28–32.
33. Cropper M, Aydede S, Portney P. Preferences for life saving programs: How the public discounts time and age. J Risk Uncertain. 1994;8:243–65.
34. Johannesson M, Johansson PO, Johanesson M, Johanesson P. Is the valuation of a QALY gained independent of age? Some empirical evidence. J Health Econ. 1997;16(5):589–99.

35. Rodríguez E, Pinto JL. The social value of health programmes: Is age a relevant factor? Health Econ. 2000;9(7):611–21.
36. Tsuchiya A. The Value of Health at Different Ages. Discussion Paper 184. York: University of York; 2000.
37. Ubel PA, Jepson C, Baron J, Mohr T, McMorrow S, Asch DA. Allocation of transplantable organs: Do people want to punish patients for causing their illness? Liver Transplant. 2001;7(7):600–7.
38. Cappelen AW, Norheim OF. Responsibility, fairness and rationing in health care. Health Policy. 2006;76(3):312–9.
39. Dolan P, Tsuchiya A. The social welfare function and individual responsibility: Some theoretical issues and empirical evidence from health. Discussion Paper 03/01. Sheffield; 2003.
40. Ottersen T. Lifetime QALY prioritarianism in priority setting. J Med Ethics. 2013;39(3):175–80.
41. Bognar G. Fair innings. Bioethics. 2015;29(4):251–61.
42. Bognar G. Priority setting and age. In: Prioritization in medicine: an international dialogue. Switserland: Springer International Publishing; 2016. 163–177 p.

Appendix 6.1: Example of reimbursement scenario

General introduction:

Healthcare resources are scarce and, therefore, health policymakers have to decide which treatments can be reimbursement from the public health insurance package, and which cannot. These decisions are often difficult, as reimbursing treatment for one patient group implies not being able to reimburse (the best possible) treatment for another.

Presented below are some of these difficult decisions. Which decision would you advise health policymakers to make in these scenarios? Which choice do you consider best in order to allocate healthcare resources as optimal as possible?

Scenario 1:

Introduction. There are two possible treatments for patients with a specific type of cancer. Only one of these treatments can be reimbursed. Without treatment, the patients have a life expectancy of 3 months and their quality of life is 3 on a scale from 0 to 10. The '0' represents the worst health possible and '10' represents the best health possible.

Question 1:

Which treatment would you advise health policymakers to choose?

- ☐ Treatment A improves patients' quality of life with 3 points (from 3 to 6), but does not influence their life expectancy.
- ☐ Treatment B improves patients' life expectancy with 3 months (from 3 to 6), but does not influence their quality of life.
- ☐ I do not have a preference for treatment A or treatment B.

Question 2:

When choosing treatment A:

Suppose that the improvement in quality of life by treatment A is unsatisfactory and less than 3 points. According to you, at which point will treatment A (improvement in quality of life) be equally good as treatment B (improvement in life expectancy of 3 months)?

- ☐ An improvement in quality of life of 2.5 points is equally good as an improvement in life expectancy of 3 months.
- ☐ An improvement in quality of life of 2.0 points is equally good as an improvement in life expectancy of 3 months.
- ☐ An improvement in quality of life of 1.5 point is equally good as an improvement in life expectancy of 3 months.

- ☐ An improvement in quality of life of 1.0 point is equally good as an improvement in life expectancy of 3 months.
- ☐ An improvement in quality of life of 0.5 point is equally good as an improvement in life expectancy of 3 months.
- ☐ An improvement in quality of life of 0 points is equally good as an improvement in life expectancy of 3 months.

When choosing treatment B:

Suppose that the improvement in life expectancy by treatment B is unsatisfactory and less than 3 months. According to you, at which point will treatment B (improvement in life expectancy) be equally good as treatment A ((improvement in quality of life of 3 points)?

- ☐ An improvement in life expectancy of 2.5 months is equally good as an improvement in quality of life of 3 points.
- ☐ An improvement in life expectancy of 2.0 months is equally good as an improvement in quality of life of 3 points.
- ☐ An improvement in life expectancy of 1.5 month is equally good as an improvement in quality of life of 3 points.
- ☐ An improvement in life expectancy of 1.0 month is equally good as an improvement in quality of life of 3 points.
- ☐ An improvement in life expectancy of 0.5 month is equally good as an improvement in quality of life of 3 points.
- ☐ An improvement in life expectancy of 0 months is equally good as an improvement in quality of life of 3 points.

When choosing the opt-out:

Why do you not have a preference for treatment A or treatment B?

- ☐ Both treatments are equally effective for these patient groups.
- ☐ Both treatments are equally ineffective for these patients groups.
- ☐ Other: …

6

Chapter 7
How does participating in a deliberative citizens panel on healthcare priority setting influence the views of participants?

Based on: Reckers-Droog VT, Jansen M, Bijlmakers L, Baltussen R, Brouwer WBF, van Exel NJA. How does participating in a deliberative citizens panel on healthcare priority setting influence the views of participants? Health Policy. 2020;124(2):143-151.

Abstract

A deliberative citizens panel was held to obtain insight into criteria considered relevant for healthcare priority setting in the Netherlands. Our aim was to examine whether and how panel participation influenced participants' views on this topic. Participants (n=24) deliberated on eight reimbursement cases in September and October, 2017. Using Q methodology, we identified three distinct viewpoints before (T0) and after (T1) panel participation. At T0, viewpoint 1 emphasised that access to healthcare is a right and that prioritisation should be based solely on patients' needs. Viewpoint 2 acknowledged scarcity of resources and emphasised the importance of treatment-related health gains. Viewpoint 3 focused on helping those in need, favouring younger patients, patients with a family, and treating diseases that heavily burden the families of patients. At T1, viewpoint 1 had become less opposed to prioritisation and more considerate of costs. Viewpoint 2 supported out-of-pocket payments more strongly. A new viewpoint 3 emerged that emphasised the importance of cost-effectiveness and that prioritisation should consider patient characteristics, such as their age. Participants' views partly remained stable, specifically regarding equal access and prioritisation based on need and health gains. Notable changes concerned increased support for prioritisation, consideration of costs, and cost-effectiveness. Further research into the effects of deliberative methods is required to better understand how they may contribute to the legitimacy of and public support for allocation decisions in healthcare.

7.1 Introduction

Priority setting in the allocation of healthcare resources is inevitable due to the increasing demand for healthcare and resulting pressure on limited budgets. Different principles have been proposed for informing allocation decisions, including the principles of maximising health and prioritising those who are worse off in terms of health [1,2]. The proposed principles to some extent all reflect a shared understanding of distributive justice; however, none addresses completely the complex and value-laden problems that arise from the need to set priorities [1,3–5]. For example, some have argued that these principles insufficiently reflect public views and preferences concerning the allocation of scarce resources [6–9]. A considerable part of the public even opposes priority setting altogether and considers access to healthcare a right to which patients are entitled without exception or restriction [10–12]. Those who do support priority setting hold different, sometimes conflicting views about the criteria that should be taken into account when setting priorities [6–9]. This heterogeneity of public views may partly explain why the outcomes of allocation decisions at times lead to public debate and controversy [12].

In a time when the public demands greater transparency and accountability from their governments and increasingly seeks opportunities to actively participate in shaping the policies that affect their lives [13], it has been argued that allocation decisions in healthcare could be improved by considering preferences from the public that are evidence-informed and elicited by means of rational democratic deliberations [3,13–15]. Such deliberative methods aim to meet the demand for a fair, legitimate, and publicly transparent way of decision making and may increase support for the outcomes of such decisions as they are more informed [3,15–17]. Examples of deliberative methods include deliberative focus groups, citizens juries, and citizens panels [17–20] that all share the following characteristics: (i) the formation of a small group of citizens who represent a larger population based on predefined characteristics, (ii) one or more meetings about the issue of interest, (iii) the preparation and dissemination of background information concerning the issue of interest, (iv) the involvement of experts to either inform the citizens or answer their questions about the issue of interest, and (v) the formulation of a set of recommendations or proposals based on the participants' deliberations [17].

Deliberative methods are increasingly applied to inform allocation decisions in healthcare, even though they are generally more time-consuming, labour-intensive, and expensive than non-deliberative methods (e.g. preference elicitation by means of surveys) [21], and very little is known about their effect. For example, empirical evidence concerning their effect on allocation decisions and the views and preferences of participants is scarce [17–22].

In the autumn of 2017, a deliberative citizens panel was held to obtain insight into participants' views and preferences concerning healthcare priority setting and identify the criteria they considered relevant for decisions concerning the composition of the basic benefits package of the health-insurance system in the Netherlands [23,24]. Health insurance is mandatory for all inhabitants of the Netherlands and the basic benefits package covers a broad range of curative and preventive treatments to protect citizens against catastrophic healthcare spending. Although in some countries deliberative citizens panels are more frequently applied, e.g. the citizens council applied by the National Institute for Health and Care Excellence (NICE) in the United Kingdom (UK) [25], this panel was the first to be applied in the Netherlands in the context of healthcare priority setting. A detailed description of the applied deliberative approach and results of the panel can be found in Bijlmakers et al. [24]. The aim of the current study was to examine whether and how panel participation influenced participants' views on healthcare priority setting. To meet this aim, we used Q methodology to investigate the views among participants before and after they participated in the panel. This methodology is increasingly applied in health services research [7,26] and to identify and describe public views on healthcare priority setting [e.g. 17,52,169,170]. In the current study, we extended previous applications of this methodology to examine changes in participants' views over time. The application of Q methodology enabled us to combine aspects of quantitative and qualitative methods to systematically examine whether and how views changed at the group level as well as the extent to which individual participants still identified with their initial viewpoints after the panel. With this study, we aim to contribute to the existing literature on the effect of applying deliberative methods for informing allocation decisions in healthcare. The approach and results of this study may be of interest to public authorities and organisations in the healthcare sector as well as in other sectors that apply, or consider applying, deliberative methods in the context of policy development and evaluation. Furthermore, the results of this study provide insight into the possible additional value of applying deliberative methods in the context of healthcare priority setting.

7.2 Methods

7.2.1 Citizens panel

Twenty-four citizens were recruited for panel participation by Motivaction; an independent research and consultancy agency in the Netherlands. The sampling was aimed at composing a varied, yet balanced, panel regarding age, sex, geographical spread, and eight mentality groups (see Table 7.1). Each of these groups represents a different set of shared values regarding work, leisure, and politics and has a distinct lifestyle and consumption pattern [24,27]. For more information on the recruitment of participants and a des-

Table 7.1 Panel characteristics (n=24)[a]

	n (%)	Mean (SD)	Min	Max
Age (Years)		44.5 (17.4)	20.0	72.0
Sex (Female)	12 (50.0)			
Education level[b]				
Middle	9 (37.5)			
High	15 (62.5)			

[a] Participants were distributed equally across the eight mentality groups and, therefore, this characteristic is omitted from the table; [b] Middle = middle vocational and secondary school, High = higher vocational and academic education.

cription of the mentality groups, we refer the interested reader to Bijlmakers et al. [24] and Motivaction [27].

The panel met during three full weekends between 16 September and 29 October 29 2017. Two experienced moderators, who were employed by Motivaction, lead the panel's deliberations on eight reimbursement cases: dental (orthodontic) braces for children, medicines and informal care for patients with Alzheimer's disease, for patients with heartburn (pyrosis), and for children with Attention Deficit Hyperactivity Disorder, the orphan drug eculizumab for patients with atypical Hemolytic Uremic Syndrome (aHUS), a total body scan, bariatric surgery and secondary prevention for patients with obesity, and a hip replacement for elderly patients [24]. These cases concern a broad range of health technologies and patient populations, and were selected to represent the variety of criteria, arguments, dilemmas, and societal values that the panel could deem relevant for setting priorities [23,24]. The first four cases were discussed during the first weekend and the latter four during the second weekend. Each case was introduced with a short video in which information was provided about the prevalence, symptoms, and course of the disease as well as the available treatment options. After watching the video, participants read written case descriptions individually and deliberated on them in small groups, followed by plenary deliberations. During the third weekend, participants were asked to prioritise the eight cases for reimbursement and discuss the trade-off between the criteria they deemed relevant for setting these priorities. In three separate plenary sessions that were held during the second and third weekend, participants were given the opportunity to discuss their questions about medical, ethical, and economic aspects of healthcare priority setting with three experts on these topics who also had expertise on the reimbursement process in the Netherlands. These experts were instructed to answer participants' questions based on their professional knowledge and experience, but not divulge their personal views on this topic. More information on the selection of the reimbursement cases and a detailed overview of the programme of the panel can be found in Bijlmakers et al. [24].

On 4 September, i.e. two weeks before the panel commenced, an information meeting was held during which the participants received general information

about the topic, objective, and procedure of the panel. The provided information was kept sparse to avoid influencing the participants before the start of the panel [24]. After the first weekend, the participants received a brochure with information about increasing healthcare expenditures, the organisation and financial structure of the Dutch healthcare system, and how healthcare priorities are currently set in the Netherlands. This information was provided to facilitate more in-depth deliberations during the second and third weekend of the panel [24].

7.2.2 Approach

Our study was conducted in three consecutive steps common to Q methodology studies [26]. First, we developed a comprehensive set of statements relating to the topic of healthcare priority setting in the Netherlands. Second, we collected data by administering the same statement-ranking exercise twice: before the participants received the information package during the information meeting and directly after the final panel meeting. Third, we analysed the collected data to examine possible changes in participants' views during the course of the panel. We describe the steps in more detail below.

7.2.3 Statement set

We developed a structured statement set that was broadly representative of our topic of interest, and hence aimed to cover all issues that participants could deem relevant for healthcare priority setting in the Netherlands. For this, we adopted the conceptual framework of the most recently conducted Q methodology study on healthcare priority setting in the Netherlands [9]. This study focused specifically on prioritising end-of-life care in the Netherlands and its framework distinguished 20 characteristics that are categorised into six domains: characteristics of the patient, characteristics of the illness, characteristics of the treatment, health effects of treatment, broader effects of treatment, and moral principles. To better align this framework with our—more general—topic of interest, we additionally inspected the framework of a Q methodology study that focused more generally on healthcare priority setting in ten European countries, among which the Netherlands [7]. After considering the relevance of the characteristics included in these two frameworks for the current study, we removed statements concerning 'prior health consumption/previous health profile', 'distribution of fixed health gains/threshold effect', and 'capacity to benefit' from the first framework [9] and included statements concerning 'rarity of the disease', 'costs/budget impact of the treatment', and 'supplier-induced demand' from the second framework [7]. We then selected 25 statements from the first framework [9] and one statement from the second framework [7], and supplemented these with two statements from related Q methodology studies that were conducted in the UK [6,8]. In order to achieve a balanced statement set that covered all issues of interest to this study, we formulated seven additional statements based on criteria and

considerations that policy makers in the Netherlands deem relevant in allocation decisions that were not yet reflected in the statement set [23,24,28,29]. Finally, we translated the statements into the Dutch language. Because the set was based on four previous carefully designed and piloted studies, no pilot test was conducted.

Table 7.3 in the results section includes the final set of 35 statements and their origin. The 20 characteristics in six domains and the associated statement numbers are presented in Appendix 7.1.

7.2.4 Data collection

All 24 participants in the citizens panel also participated in this study. This sample size was sufficient for the purpose of this analysis [26,30]. The participants completed the first statement-ranking exercise directly after the information meeting on September 4, 2017 (T0) and the second during the final panel meeting on October 29, 2017 (T1). Before performing the exercise, participants received an oral group instruction on how to perform the exercise from one of the researchers (MJ). They received a copy of these instructions on paper (see Appendix 7.2), for reference. This researcher remained present during the exercise in case participants had any questions about the procedure. Subsequently, participants received a set of the 35 statements printed on cards, a sorting grid (see Appendix 7.3), and a response sheet. Participants first read all statements and divided them into three piles ('agree', 'disagree', and 'neutral'). Then, they re-read the statements in the 'agree' pile, selected the two they agreed with most, and placed them in column 9 of the sorting grid, followed by placing the next three statements they then agreed with most in column 8 and so on until they finished this pile. Next, they followed the same procedure for the 'disagree' pile, starting with column 1, and finally placed the statements in the 'neutral' pile in the remaining open spots in the middle of the grid. After finishing the exercise, participants used the response sheet to explain in writing their motivation for placing the statements in the extreme positions of the grid, i.e. columns 1 and 9. The columns were presented to participants as being from 1 to 9 on the sorting grid to avoid imposing connotations of negative, neutral or positive to columns of the grid; however, we recoded the columns to -4 to +4 for the analysis of the data and interpretation of the viewpoints.

7.2.5 Data analysis

We conducted a principal component analysis followed by oblimin rotation to identify groups of participants with highly (Pearson) correlated statement rankings at both time points separately. This type of oblique rotation method is typically used to allow for a non-orthogonal rotation. We selected the best number of factors from all possible factor solutions that were supported by the data by applying the criteria: (i) eigenvalues of factors >1 and (ii) a

minimum of two non-confounded 'exemplars' per factor. Exemplars are those participants with (i) a factor loading above the significance threshold of 0.33 ($p<0.05$; calculated as $1.96/\sqrt{35}$, where 35 is the number of statements) and (ii) for whom the square of the loading for a factor is larger than the sum of the square loadings for all other factors [30,31]. Based on inspections of the correlations between factors and the interpretation of the factors in each factor solution, we selected the factor solution that lead to the most intelligible reduction of the data. Subsequently, we computed factor arrays for each factor. These arrays represent how a participant with a correlation of 1 with a factor would have ranked the statements. We used the factor arrays, including the characterising and distinguishing statements, for interpreting the factors as viewpoints. Characterising statements are those that hold the positions -4, -3, +3, and +4 in the factor arrays, and as such represent the statements that participants with a specific viewpoint least and most agreed with. Distinguishing statements are those with a statistically significantly different position in a factor array from their position in the array of at least one other factor ($p<0.05$; calculated based on the absolute difference in z-scores of statements between the factor arrays). We used the verbatim quotes of exemplars that we obtained from the response sheets to help describe the viewpoints in the wording of the participants.

We examined changes in viewpoints in multiple ways. At the level of the viewpoints, we examined the correlations and the main similarities and differences between the viewpoints at T0 and T1. At the level of the participants, we examined the extent to which participants associated themselves with the initial viewpoints, i.e. the viewpoints identified at T0, after they participated in the panel (at T1). For this, we combined the data of T1 with the factor arrays of T0 and calculated the mean (SD) difference in correlation with the initial viewpoints between T0 and T1. Furthermore, we examined the transitions between viewpoints made by exemplars over time and the extent to which the views of participants converged after the panel. We did this by examining the mean (SD) correlations of the statement rankings between participants at T0 and T1 and applying an F-test for small sample sizes to examine the difference in the associated variances.

We used Cohen's classification system for interpreting the obtained correlation coefficients [32]. In line with this system, we interpreted correlations below 0.30 as low, between 0.30 and 0.50 as moderate, and above 0.50 as high.

We used the qmethod package in Rstudio 1.0.143 (Rstudio, Inc., Boston, MA, USA) for conducting the analyses [31].

7.2.6 Ethics

The Committee on Research Involving Human Subjects of the Radboud University Medical Center reviewed and waived ethical approval for this study (reference 2017-3444).

7.3 Results

Table 7.2 presents the participants' factor loadings with the viewpoints at T0 and T1, respectively.

Table 7.2 Factor loadings at T_0 and T_1 (n=24)

	Views at T_0			Views at T_1		
id	1	2	3	1	2	3
1	0.74*	0.02	-0.28	0.85*	-0.17	0.10
2	0.72*	0.00	0.19	0.32	0.37	-0.33
3	0.50	0.55*	-0.06	0.16	0.67*	-0.35
4	0.58*	0.46	0.25	0.80*	0.14	0.27
5	0.94*	-0.20	-0.01	0.66*	0.14	-0.39
6	0.37	0.28	0.28	0.03	0.69*	-0.23
7	0.77*	0.21	-0.15	0.42	0.47	0.25
8	0.31	0.28	-0.52*	0.21	0.49*	0.10
9	-0.13	0.03	0.71*	0.12	0.14	0.54*
10	-0.01	-0.01	0.82*	0.14	-0.13	-0.22
11	-0.16	0.89*	0.10	0.13	0.73*	0.02
12	-0.09	0.79*	-0.18	-0.38	0.85*	0.01
13	0.81*	-0.04	0.21	0.62*	-0.03	-0.37
14	0.62*	-0.31	0.39	0.30	0.11	-0.73*
15	0.20	0.76*	0.02	0.59*	0.42	0.13
16	0.91*	-0.01	-0.17	0.60*	0.36	-0.01
17	0.75*	-0.13	0.04	0.40	0.05	0.63*
18	0.84*	0.06	0.00	0.86*	-0.08	-0.22
19	0.47	0.63*	0.07	0.18	0.83*	-0.03
20	0.25	0.48*	0.11	-0.02	0.78*	-0.03
21	-0.28	0.78*	-0.25	-0.13	0.73*	0.38
22	0.78*	-0.01	-0.14	0.85*	-0.05	0.03
23	0.05	0.68*	0.01	0.15	0.54	0.53
24	0.72*	0.22	0.13	0.47	0.26	-0.42
Explained variance (%)	33.6	19.2	8.8	22.6	22.5	11.1
Exemplars[a] (n)	12	8	3	8	8	3

[a] The factor loadings of exemplars are indicated with an asterisk (*). These loadings meet the following two criteria: (i) the loading is above the significance threshold of 0.33 ($p<0.05$, calculated as $1.96/\sqrt{35}$, where 35 is the number of statements) and (ii) the square of the loading for a factor is larger than the sum of the square loadings for all other factors [26,30].

Table 7.3 Factor arrays at T_0 and T_1

		Views at T_0			Views at T_1		
#	Statement	1	2	3	1	2	3
1	Younger people should be given priority over older people, because they haven't had their fair share of health yet.[a]	-2*	-3	+1	-2	-3	+2*
2	Children's health should be given priority over adults' health.[a]	-2*	-2	+4	-1*	-3*	+2*
3	Individual responsibility should not be taken into account because people don't always have control over their way of living.[a]	+1*	0*	-4*	+1*	-2	-2
4	The health system should be about looking after those patients in greatest need.[a]	0	0	-1	+1	0	+1
5	Priority should be given to those treatments that generate the most health.[b]	0*	+2*	-2*	+1	0	+3*
6	Priority should be given to restoring health to a level that is sufficient for people to participate in their usual activities.[a]	+1	+2	+1	+1	+2	0*
7	Priority should be given to preventive healthcare.[c]	+1	+4*	0	+2	+1	0*
8	Patients with a family should be prioritised because their treatments will benefit others as well as the patient themselves.[a]	-1	-2	+3*	-4	-4	0*
9	Treatments that are very costly in relation to their health benefits should not be reimbursed.[a]	-2	+1*	-1	-2*	-1*	+3*
10	The health system should restrict itself to treatments that have proven to bring about health gains.[e]	-1*	+1*	-4*	0	+1	+3*
11	We should support patients' choice for treatment, even if it is very costly in relation to its health benefits.[a]	+2*	-2	-2	0*	-2	-2
12	It's important to respect the wishes of patients who feel they should take every opportunity to extend their life.[a]	+2*	-3*	0*	+1*	-2	-2
13	Poorer people should be given priority because they don't have the same opportunities in life.[a]	-3	-4	+1*	-4*	-4*	-1*
14	If there is a way of helping patients, it is morally wrong to deny them this treatment.[a]	+3*	0	0	+4*	0*	-3*
15	If you choose to spend a lot of money on a specific patient group, you have to realise that there will be less money left for other patient groups.[e]	+1	+1	+2	+2	+3	+1
16	Access to healthcare should be based on need for care, not on patient characteristics, such as their gender, age, or ethnicity.[a]	+3	+2	0*	+4	+3	-3*
17	You can't put a price on life.[a]	+4	+1*	+4	+2*	0*	-4*

18	Priority should be given to younger people, because they may benefit from treatment for longer.[a]	-3	-4	+2*	-3	-3	+2*
19	People who live a healthy life should be prioritised over people with an unhealthy lifestyle.[a]	-4*	+1	+1	-1*	0	+1
20	People with a severe condition should be treated with priority over people with a non-severe condition.[a]	+2*	-1	-1	+2	-1*	+2
21	If a treatment is costly in relation to its health benefits, but the only treatment available, it should still be provided.[a]	+3*	-1	0	+3*	0**	-2*
22	There is no sense in saving lives if the quality of those lives will be really bad.[a]	0*	+2*	-3*	0*	+2*	+4*
23	There is no point in providing treatments that do not generate considerable health benefits.[e]	0*	+3*	-2*	0*	+1*	+4*
24	Treatment of illnesses that put a high burden on patients' families should receive priority.[a]	-1	-1	+3*	-2	-1	-1
25	Everyone has a right to healthcare, but this doesn't mean that everything can always be reimbursed.[a]	0	+3*	0	0	+3*	+1
26	At the end of life it is more important to provide a death with dignity than treatments that will only extend life for a short period of time.[a]	0*	+4*	-3*	0*	+4*	+1*
27	The health system should be about getting the greatest health benefit overall for the population.[a]	+1*	+3*	-1*	-1	+2*	0
28	People who are ill through no fault of their own should receive priority over people who in some way are responsible for their own illness.[a]	-4*	0*	+2*	-2	+1*	-1
29	If it is possible to save a life, every effort should be made to do so.[a]	+4	-3*	+3	+3*	-1	0
30	People can pay for inexpensive treatments out of pocket.[e]	-1	-1	-1	0	+4*	0
31	If the total costs of treating a disease (for all patients) are high, this treatment should receive less priority.[e]	-3	-1	-2	-1	-1	0
32	A treatment may cost more if it is not only beneficial for the patient but also for society.[e]	-1*	0*	+2*	-3	+1*	-3
33	Priority should be given to people with rare diseases, even when these diseases do not necessarily cause more health damage than more common ones.[b]	0*	-2	-3	-3	-2	-1
34	Medical tests for the early detection of diseases that often lead to unnecessary treatments should not be reimbursed.[e]	-2*	0	0	-1*	+2*	-4*
35	If a treatment is the only available treatment for a disease, it should be reimbursed.[d]	+2*	0	+1	+3*	0*	-2*

[a] Statement from Wouters et al. [9]; [c] Statement from van Exel et al. [7]; [d] Statement from Baker et al. [6]; [e] Statement from McHugh et al. [8]; [f] Based on considerations that the Dutch National Health Care Institute deemed relevant in the reimbursement cases; * Distinguishing statement, i.e. statement with a statistically significantly different position in a factor array from their position in the array of at least one other factor ($p<0.05$).

Table 7.3 on the previous pages presents the factor arrays, including the characterising and distinguishing statements. Below, we describe the viewpoints before and after the panel and discuss the changes in viewpoints. We present the numbers of the most relevant statements within parenthesis with a hashtag (#), followed by their position in a factor array, e.g. (#1,+4). Distinguishing statements are presented with an additional asterisk, e.g. (#1,+4*). Verbatim quotes of exemplars are presented within quotation marks, followed by their identification number, e.g. (id2).

7.3.1 Viewpoints before the panel

At T0, we identified three factors that together explained 61.6% of the variance in the statement rankings. The correlations between the factors were low to moderate (ρ=0.14 for 1 vs. 2, ρ=-0.01 for 1 vs. 3, ρ=-0.30 for 2 vs. 3). The factors had 12, eight, and three exemplars, respectively. Factor 3 had two positive exemplars (id9 and id10) and one negative exemplar (id8) and was, therefore, interpreted as being bipolar. One participant (id6) was statistically significantly associated with factor 1; however, did not meet the second criterion for being identified as an exemplar.

Viewpoint 1

People with viewpoint 1 considered access to healthcare as a right and believed that everyone should have equal access to healthcare. According to people with this view access should solely be based on patients' need for care and not on their personal characteristics, such as their gender, age, ethnicity (#16,+3; #18,-3), lifestyle (#19,-4*; #28,-4*), or socio-economic status (#13,-3). "Everyone has a right to healthcare [and] personal characteristics are not important at all" (id7). As prioritisation in healthcare should be based on patients' need for care, "there should be no discrimination [between patients]" (id24). People with this view believed that healthcare costs should play no role in priority setting as "you cannot regard a life in an economic way" (id4). If there is a way of helping patients, it is morally wrong to deny them this treatment (#14,+3*). People holding this view did not believe that a treatment should receive less priority if the total costs of treating a disease (for all patients) are high (#31,-3). Rather, if a treatment is costly in relation to its benefits, but is the only treatment available, it should still be provided (#21,+3*). People with this view also believed that patients' choice for treatment should be supported, even if it is very costly in relation to its benefits (#11,+2*). "Everyone has a right to healthcare; even when there is no or little treatment benefit you cannot deny treatment [to patients]!" (id 13). They emphasised that you cannot put a price on life (#17,+4) and if it is possible to save a life, every effort should be made to do so (#29,+4). "Regardless of money, if it is possible, a life has to be saved" (id18).

Viewpoint 2

People with viewpoint 2 believed that everyone has a right to healthcare, but that this does not mean that everything can always be reimbursed (#25,+3*). "Everyone is insured and has […] a right to healthcare, but not everything can always be covered by the [public health] insurance" (id15). As "healthcare costs keep rising, there should be restricting measures" (id21). People with this view emphasised the importance of the effectiveness of treatments. The health system should be about getting the greatest benefit overall for society (#27,+3*) and there is no point in providing treatments that do not generate considerable health benefits (#23,+3*). Accordingly, they support prioritisation based on treatment characteristics, such as the type and size of health gains from treatment, but like viewpoint 1, they oppose prioritisation based on patient characteristics, such as their age (#18,-4). They further emphasised that, at the end of life, it is more important to provide a death with dignity than treatments that may extend life only for a short period of time (#26,+4*). They neither believed that, if it is possible to save a life, every effort should be made to do so (#29,-3*) nor that it is important to respect the wishes of patients who feel they should take every opportunity to extend their life (#12,-3*). They do believe that priority should be given to preventive healthcare (#7,+4*), because "this can save a lot of money" (id12).

Viewpoint 3

People with viewpoint 3 were positively oriented towards prioritisation based on patient characteristics, such as their age. They believed that children should be given priority over adults (#2,+4), because they may benefit from treatment longer (#18,+2*). "Children hold the future and, if […] a choice has to be made, the child is the first one entitled to receiving care" (id9). However, they opposed prioritisation based on lifestyle (#3,-4*). People holding this view also found that broader treatment effects should be taken into consideration. They believed that treatment of illnesses that put a high burden on families of patients should receive priority (#24,+3*), because treating these patients benefits them as well as others (#8,+3*). Consequently, treatments that are beneficial for both the patient and society should be allowed to cost more (#32,+2*). Although being positively oriented towards prioritisation in healthcare, they emphasised that you cannot put a price on life (#17,+4) and that, if it is possible to save a life, every effort should be made to do so (#29,+3). They believed there is a sense in saving lives, even if the quality of those lives will be really bad (#22,-3*), and in providing treatments that do not generate considerable health gains (#23,-2*).

In contrast, people who opposed this viewpoint were in favour of priority setting based on lifestyle (#3,-4*). They also believed that priority should be given to those treatments that generate the most health #5,-2*) and that the health system should restrict itself to treatments that have proven to bring

about health gains (#10,-4*). "If there is evidence that a treatment is effective, it should always be reimbursed" (id8).

7.3.2 Viewpoints after the panel

At T1, we identified three factors that together explained 56.3% of the variance. The correlations between viewpoints were again low to moderate (ρ=0.30 for 1 vs. 2, ρ=-0.18 for 1 vs. 3, ρ=0.06 for 2 vs. 3). The factors had eight, eight, and three exemplars, respectively. Factor 3 had two positive exemplars (id9 and id17) and one negative exemplar (id14) and was, therefore, interpreted as being bipolar. Four participants (id2, id7, id23, and id24) were 'mixed loaders' as they were statistically significantly associated with more than one factor. They did not meet the second criterion for being identified as exemplars. One participant (id10) was a 'null loader' as s/he was not statistically significantly associated with any of the factors.

Factors 1 and 2 at T1 strongly resembled factors 1 and 2 at T0, with ρ=0.84 and ρ=0.78, and hence can be regarded as slightly different manifestations of their corresponding viewpoints at T0. Therefore, we describe only the main similarities and differences between these viewpoints at T0 and T1. The correlation between factors 3 at T0 and T1 was ρ=0.32 and, therefore, we regard and describe factor 3 at T1 as a newly emerged viewpoint.

Viewpoint 1

Before the panel, people with viewpoint 1 emphasised equal access to care and that all treatments should be available for patients. Like people with this view before the panel, people with viewpoint 1 at T1 believed that it is morally wrong to deny patients treatment, if there is a way of helping them (#14,+4*) or if a treatment is the only one available (#35,+3*). They also believed that access to care should be based on need and not on patient characteristics, such as their gender, age, ethnicity (#16,+4), or or socio-economic status (#13,-4*). However, people with this view less were strongly opposed to prioritisation based on lifestyle than those with viewpoint 1 at T0 (#19,-1*; #28,-2) and more strongly opposed to prioritisation based on characteristics of the illness, such as its rarity (#33,-3). They were notably more considerate of treatment costs. They believed less strongly that you cannot put a price on life (#17,+2*) and that treatment should always be supported, even if it is very costly in relation to its health benefits (#11,0*). They also believed less strongly that a treatment may cost more if it is not only beneficial for a patient but also for society (#32,-3).

Viewpoint 2

Before the panel, people with viewpoint 2 believed that everyone has an equal right to healthcare and emphasised the importance of treatment effectiveness and efficiency. Like people with this view before the panel, people with vie-

wpoint 2 at T1 believed that everyone has a right to healthcare, but that this does not mean that everything can always be reimbursed (#25,+3*). "There simply is a limited budget [and] choices have to be made" (id19). People with this view believed that access to care should be based on need for care and not on patient characteristics, such as their gender, age, ethnicity (#16,+3; +18,-3), or socio-economic status (#13,-4*). However, people with this view were less strongly opposed to prioritisation based on lifestyle (#3,-2). They believed more strongly than those with viewpoint 2 at T0 that inexpensive treatments can be paid out of pocket (#30,+4) as "it is relatively cheap" (id6) and "does not really affect patients' disposable income" (id12). They also believed more strongly that medical tests for the early detection of diseases that often lead to unnecessary treatments, should not be reimbursed (#34,+2*) and that if you choose to spend a lot of money on a specific patient group, you have to realise there will be less money left for other groups (#15,+3). For people with this viewpoint, it was "more important that patients can die with dignity" (id20) than to extend life for a short period of time (#26,+4*).

Viewpoint 3

People with viewpoint 3 at T1 believed that prioritisation should be based on the health effect of treatment and patient characteristics such as their gender, age, and ethnicity (#1,+2*; #2,+2*; #13,-1*; #16,-3*; #18,+2*), and lifestyle (#3,-2; #19,+1; #28,-1). "People do have control over their lives, they cannot live recklessly and still benefit" (id9). According to people with this view, the health system should restrict itself to treatments that have proven to bring about health gains (#10,+3*). They considered treatments that generate the most health to be the most important (#5,+3*) and believed there is neither a point in providing treatments that do not generate significant health gains (#23,+4*), nor in saving lives if the quality of those lives will be really bad (#22,+4*). People with this view did not agree with the statements that you cannot put a price on life (#17,-4*) and that it is morally wrong to deny patients treatment (#14,-3). They believed that treatments that are very costly in relation to their health gain should not be reimbursed (#9,+3*). Nonetheless, they disagreed that medical tests for the early detection of diseases, that often lead to unnecessary treatments, should not be reimbursed (#34,-4*).

In contrast, people who opposed this viewpoint believed that "costs are not the only thing that matters" (id14). If a treatment is the only available treatment for a disease it should be reimbursed and if it is not only beneficial for the patient but also for society it may cost more (#32,-3; #35,-2*). "If costs need to be taken into account, people can pay for inexpensive treatments themselves in order to reimburse expensive treatments [from public funding]" (id14).

Table 7.4 Factor loadings on the initial viewpoints (i.e. those identified at T_0) before (at T_0) and after (at T_1) the panel and the difference in factor loadings between the two time points (n=24)

	View 1 at T_0			View 2 at T_0			View 3 at T_0		
id	T_0[a]	T_1	T_1 - T_0	T_0[a]	T_1	T_1 - T_0	T_0[a]	T_1	T_1 - T_0
1	0.74	0.73	-0.02	0.02	0.24	0.23	-0.28	-0.01	0.27
2	0.72	0.42	-0.30	0.00	0.12	0.12	0.19	0.04	-0.15
3	0.50	0.35	-0.15	0.55	0.57	0.02	-0.06	-0.11	-0.05
4	0.58	0.49	-0.09	0.46	0.28	-0.18	0.25	0.01	-0.25
5	0.94	0.80	-0.14	-0.20	0.22	0.42	-0.01	-0.10	-0.09
6	0.37	0.29	-0.08	0.28	0.49	0.22	0.28	-0.16	-0.44
7	0.77	0.37	-0.41	0.21	0.51	0.30	-0.15	-0.42	-0.27
8	0.31	0.28	-0.03	0.28	0.44	0.17	-0.52	-0.18	0.34
9	-0.13	-0.09	0.04	0.03	0.11	0.08	0.71	0.11	-0.59
10	-0.01	0.11	0.12	-0.01	-0.21	-0.20	0.82	0.77	-0.05
11	-0.16	0.26	0.42	0.89	0.59	-0.30	0.10	-0.04	-0.13
12	-0.09	-0.28	-0.19	0.79	0.46	-0.33	-0.18	-0.25	-0.07
13	0.81	0.73	-0.08	-0.04	0.03	0.07	0.21	0.14	-0.06
14	0.62	0.54	-0.08	-0.31	-0.13	0.18	0.39	0.40	0.01
15	0.20	0.59	0.39	0.76	0.53	-0.23	0.02	-0.23	-0.24
16	0.91	0.56	-0.34	-0.01	0.28	0.29	-0.17	-0.31	-0.14
17	0.75	0.19	-0.55	-0.13	0.14	0.28	0.04	-0.22	-0.26
18	0.84	0.71	-0.13	0.06	0.08	0.02	0.00	0.04	0.04
19	0.47	0.31	-0.16	0.63	0.71	0.08	0.07	-0.26	-0.33
20	0.25	0.19	-0.06	0.48	0.70	0.22	0.11	-0.27	-0.38
21	-0.28	-0.15	0.13	0.78	0.59	-0.19	-0.25	-0.24	0.01
22	0.78	0.68	-0.11	-0.01	0.24	0.24	-0.14	-0.12	0.02
23	0.05	0.04	-0.02	0.68	0.79	0.11	0.01	-0.28	-0.29
24	0.72	0.57	-0.16	0.22	0.19	-0.02	0.13	-0.34	-0.47
Mean (SD) difference	NA	NA	-0.08 (0.21)	NA	NA	0.07 (0.21)	NA	NA	-0.15 (0.22)

NA, Not Applicable; [a] These factor loadings correspond with the factor loadings at T_0 presented in Table 7.2.

7.3.3 Association with initial viewpoints

The mean (SD) correlation between participants' statement rankings at T0 and T1 was 0.57 (0.17), ranging from 0.19 to 0.78 (see Appendix 7.4). For 18 participants the correlation between T0 and T1 was strong, for three moderate, and for another three low. Although none of the participants ranked the statements in exactly the same way, these relatively high correlations indicate that the views of most participants were largely similar before and after the panel.

Table 7.4 presents the extent to which participants associated themselves with the initial viewpoints, i.e. the viewpoints from before the panel (at T0), after they participated in the panel (at T1). These results show that most participants (n=19) correlated less strongly with the initial viewpoint 1 at T1, with a mean (SD) decrease in correlation of 0.08 (0.21). Of the participants, 17 correlated more strongly with the initial viewpoint 2 at T1 with a mean (SD) increase in correlation of 0.07 (0.21) and 18 correlated less strongly with the initial viewpoint 3 at T1, with a mean (SD) decrease in correlation of 0.15 (0.22).

Table 7.5 presents the transitions between viewpoints made by exemplars over time. These results show that of the 12 exemplars with viewpoint 1 at T0, seven made no transition and still adhered to this viewpoint, two changed their view to viewpoint 3, and three were no longer associated with one of the viewpoints at T1. Of the eight exemplars with viewpoint 2, six made no transition and still adhered to this viewpoint, one changed his/her view to viewpoint 1, and one was no longer associated with one of the viewpoints at T1. Of the three exemplars with viewpoint 3, none still adhered to this viewpoint at T1. One exemplar changed his/her view to viewpoint 2, one changed his/her view to the new viewpoint 3, and one was no longer associated with one of the viewpoints at T1.

Table 7.5 Transition matrix of exemplars' views

		Views at T_1				
		1	2	3	No distinct viewpoint	Total
Views at T_0	1	7	NA	2	3	12
	2	1	6	NA	1	8
	3	NA	1	1	1	3
	No distinct viewpoint	NA	1	NA	NA	1
	Total	8	8	3	5	24

NA, Not Applicable.

7.3.4 Convergence between views

At T0, the mean (SD) correlation between participants' statement rankings was 0.32 (0.28), ranging from -0.43 to 0.79. At T1, this was 0.32 (0.25), ranging from -0.26 to 0.73. See Appendices 7.5 and 7.6 for the correlation matrices of participants' rankings at T0 and T1. The difference in variance decreased marginally between the rankings at both time points ($p<0.001$), indicating some modest convergence between the views of participants over time.

7.4 Discussion

In this study, we examined whether and how participation in a deliberative panel influenced the views of participants on healthcare priority setting. Our main finding is that participants' views before and after the panel partly remained stable. There was a strong resemblance between two of the three views identified before and after the panel, while the third view was distinctly different at both time points and 18 participants showed high correlation between their views at T0 and T1. Equal access to healthcare, prioritisation based on patients' needs, and the relevance of the size and type of treatment benefits remained important during the course of the panel. We observed two notable changes. Firstly, support for prioritisation in healthcare generally seems to have increased after panel participation. Secondly, participants became more considerate of healthcare costs and of cost-effectiveness as a relevant criterion for setting priorities in healthcare.

To our knowledge, this study is one of the few to examine changes in views on healthcare priority setting through deliberation and the first to do so in the Netherlands. This limits us in our ability to compare our results with those of other studies. However, we can compare our results to two other studies that examined the effect of deliberation on views in the context of healthcare priority setting and two Q methodology studies that examined views on this topic in the Netherlands. Dolan et al. [21] examined the effect of deliberation on views in a sample of 60 patients in the UK. They observed a trend towards treating different patient groups more equally and participants who were initially unwilling to prioritise between patient groups remained so after deliberation. Abelson et al. [17] examined the effect of deliberation in a sample of 46 participants in Canada, by using a controlled design. They found that participants' views became more susceptible to change when more deliberation was introduced. Participants who changed their view did so in a similar direction, indicating that deliberation may lead to increased consensus among participants. Like in these studies, we found views opposing priority setting that remained relatively stable and that deliberation can lead to changes in viewpoints as well as to convergence between them. Van Exel et al. [7] and Wouters et al. [9] applied Q methodology to examine views on healthcare priority setting in the Netherlands. Like in these studies, we found that members

of the public—before deliberation—generally hold a view on priority setting that emphasises the importance of equal access and disregards costs, while some recognise the scarcity of healthcare resources and are willing to accept certain criteria for setting priorities.

Before discussing the main strengths and limitations of our study, we would like to reflect on the bipolar nature of viewpoints 3 at T0 and T1. Previous literature shows that there are different ways to deal with the computation and interpretation of bipolar factors. Some have argued that negative exemplars should be excluded from the computation of the factor array as this leads to a clearer, or purer, interpretation of the positive pole of the viewpoint [e.g. 199]. Others have argued that negative exemplars should be included in the computation of the factor array. Excluding them would lead to an unbalanced interpretation of the factor, as it no longer fully represents the views of the participants who define the factor (albeit on different sides of the pole) [26]. Here, we followed the latter argument and chose to retain the negative exemplars in the computation of the factor arrays and the interpretation of the bipolar factors 3 at T0 and T1. In order to explore the implications of this choice, we also inspected a solution excluding the negative exemplars. At T0, the correlation between factors 3 with and without negative exemplars was 0.97, and hence these factors seem to portray the same view. At T1, the correlation between factors 3 with and without negative exemplars was 0.65 and the positioning of some statements changed considerably. More specifically, compared to the interpretation presented in the Results section, the viewpoint would agree less strongly that personal characteristics should be taken into account in healthcare priority setting (#1,+1*; #2,+1*; #13,-4*; #16,-1*; #18,+1*), and more strongly that individual responsibility is relevant (#3,-3*) and inexpensive treatments can be paid out of pocket (#30,+4). Although excluding the negative exemplar leads to a slightly different viewpoint 3 at T1, it remains a new view as compared to viewpoint 3 at T0 (excluding the negative exemplar; ρ = -0.03), and, therefore, does not affect the main finding of our study.

The main strength of our study lies in the repeated use of Q methodology to examine in depth whether and how deliberation influences views on healthcare priority setting. To our knowledge, this approach has not been applied before, neither in nor outside the field of healthcare. Despite this strength, some limitations need to be discussed. Firstly, although we speak of the 'influence' of deliberation on views, no causal conclusions can be drawn in the absence of a control group. Secondly, the reimbursement cases may have primed the need for setting priorities and the relative importance of certain characteristics after the panel. We do note that the cases were carefully selected to represent all issues participants may have deemed relevant for setting priorities in a broad range of health technologies and patient populations. In that sense, they were aligned with the broad considerations represented in the statement set. Therefore, insofar the cases influenced the statement rankings after the

panel we think this influence is relevant in the context of this study. Finally, lower-educated people are not represented in the panel. However, this is only problematic if they differ from higher-educated people with respect to their susceptibility for deliberation. This we do not know and would be a relevant topic for further research.

Our study contributes to the limited literature on the effect of deliberative methods by giving insight into whether and how deliberation influences views on healthcare priority setting. Based on our results, some questions can be raised regarding the application of deliberative methods in the context of healthcare priority setting. For example, if the purpose is to inform allocation decisions, questions can be raised about the extent to which participants' views over time still represent the actual views of the public. If the latter is desired in a panel, one could argue that the time anyone participates in such a panel should be restricted and that panel participants should regularly be replaced by other members of the public. However, if changes in views, as observed here, are interpreted as the effect of learning and the purpose is that better informed and more considered views are represented in a panel, it can also be argued that panel members should participate in a panel for a longer period of time. In this case, one could also argue against the application of a deliberative citizens panel and in favour of better information provision to the public and more public debate, through which a similar learning effect perhaps can be achieved in members of the public at large. Notwithstanding, it is important to note that it is unlikely that any one of these approaches will lead to public consensus about allocation decisions. The recurrent finding in the literature that views on priority setting in healthcare differ and can conflict, together with the current finding that views remain diverse and only moderately converge after deliberation, suggests that any allocation decision will probably still be met with opposition from some group in society. Still, insight into the diversity of views is important to be able to understand the opposition that allocation decisions can bring about and how the outcomes of decisions, if so desired, can be better aligned with societal preferences.

We appreciate that, based on the design and results of the current study it remains unclear why exactly participants' views changed and the extent to which their views changed under the influence of, for example, the other participants, information provided, and experts consulted. If changes do not result from the deliberations, but rather from external influences (e.g. from stakeholders, such as experts, patients, and industry), a deliberate panel may have limited additional value as these views usually are already represented in allocation decisions. The crucial question in this context is the purpose of applying deliberative panels. Is it for policy makers to consult citizens or give them a vote in allocation decisions, strengthen the appraisal of available evidence, increase the legitimacy of decisions, or rather to predict or increase public support for the outcomes of such decisions? Regardless of the purpose, it is important that citizens contribute in a way that is complementary to

other stakeholders. Although answering these questions lies outside the scope of this paper, they are related to the issue that panel participants may experience (moderate) changes in their viewpoints over time. Moreover, they emphasise that further research is indispensable for applying these methods in a way that contributes to the legitimacy of and public support for allocation decisions in healthcare.

7.5 Conclusions

Our study showed that participants' views partly remained stable over the course of the panel, specifically regarding equal access to healthcare, prioritisation based on patients' needs, and the importance of the size and type of treatment benefits. Notable changes after deliberation concerned the increased support for prioritisation, consideration of costs, and relevance of a cost-effectiveness criterion in allocation decisions. Considering the increasing interest in deliberative methods among policy makers in healthcare and the limited empirical evidence concerning the effect of deliberative methods on participants' views and preferences, further research is required to better understand how deliberative methods can contribute to the legitimacy of and public support for the outcomes of allocation decisions in healthcare.

References

1. 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):1–10.
2. Hernæs UJV, Johansson KA, Ottersen T, Norheim OF. Distribution-Weighted Cost-Effectiveness Analysis Using Lifetime Health Loss. Pharmacoeconomics. 2017;35(9):965–74.
3. Fleck L. Just Caring: In Defense of the Role of Democratic Deliberation in Health Care Rationing and Priority-Setting. In: Den Exter A, Buijsen M, editors. Rationing Health Care Hard Choices and Unavoidable Trade-Offs. Antwerpen: Maklu; 2012. p. 25–45.
4. 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 Health Econ. 2013;14(1):107–15.
5. 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):621–9.
6. Baker R, Wildman J, Mason H, Donaldson C. Q-Ing For Health—A New Approach To Eliciting The Public's Views On Health Care Resource Allocation. Health Econ. 2014;23:283–97.
7. 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;126:128–37.
8. McHugh N, Baker RM, Mason H, Williamson L, Van Exel J, Deogaonkar R, et al. Extending life for people with a terminal illness: A moral right and an expensive death? Exploring societal perspectives. BMC Med Ethics. 2015;16(1):1–15.
9. Wouters S, van Exel J, Baker R, B.F. Brouwer W. Priority to End of Life Treatments? Views of the Public in the Netherlands. Value Health. 2017;20(1):107–17.
10. Mason H, van Exel J, Baker R, Brouwer W, Donaldson C, Pennington M, et al. From representing views to representativeness of views: Illustrating a new (Q2S) approach in the context of health care priority setting in nine European countries. Soc Sci Med. 2016;166:205–13.
11. Mason H, Collins M, McHugh N, Godwin J, Van Exel J, Donaldson C, et al. Is "end of life" a special case? Connecting Q with survey methods to measure societal support for views on the value of life-extending treatments. Health Econ. 2018;27(5):819–31.
12. 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):1–18.
13. Caddy J. Caddy J. Evaluating public participation in policy-making. Paris: Publications de l'OCDE; 2005.
14. Daniels N. Just health: M eeting health needs fairly. Cambridge: Cambridge University Press; 2008.

15. Jansen MP, Helderman JK, Boer B, Baltussen R. Fair processes for priority setting: Putting theory into practice: Comment on "expanded HTA: Enhancing fairness and legitimacy." Int J Health Policy Manag. 2017;6(1):43–7.
16. Abelson J, Forest P-G, Eyles J, Smith P, Martin E, Gauvin F-P. Obtaining public input for Past experiences and future prospects. Can Public Adm. 2002;45(1):70–97.
17. Abelson J, Eyles J, McLeod CB, Collins P, McMullan C, Forest PG. Does deliberation make a difference? Results from a citizens panel study of health goals priority setting. Health Policy. 2003;66(1):95–106.
18. Degeling C, Carter SM, Rychetnik L. Which public and why deliberate? A scoping review of public deliberation in public health and health policy research. Soc Sci Med. 2015;131:114–21.
19. Abelson J, Wagner F, DeJean D, Boesveld S, Gauvin FP, Bean S, et al. Public and patient involvement in health technology assessment: A framework for action. Int J Technol Assess Health Care. 2016;32(4):256–64.
20. Mitton C, Smith N, Peacock S, Evoy B, Abelson J. Public participation in health care priority setting: A scoping review. Health Policy. 2009;91(3):219–28.
21. Dolan P, Cookson R, Ferguson B. Effect of discussion and deliberation on the public's views of priority setting in health care: Focus group study. BMJ. 1999;318(7188):916–9.
22. Street J, Duszynski K, Krawczyk S, Braunack-Mayer A. The use of citizens' juries in health policy decision-making: A systematic review. Soc Sci Med. 2014;109:1–9.
23. Baltussen R, Bijlmakers L, Jansen M, Team P. Draagvlak voor lastige keuzes. Eindrapport van het Burgerforum 'Keuzes in de zorg' [Internet]. 2018 [cited 2018 Jul 2]. Available from: http://hdl.handle.net/1765/114121
24. Bijlmakers L, Jansen M, Boer B, van Dijk W, Groenewoud S, Zwaap J, et al. Increasing the Legitimacy of Tough Choices in Healthcare Reimbursement: Approach and Results of a Citizen Forum in The Netherlands. Value Health. 2020;23(1):32–8.
25. National Institute for Health and Care Excellence. Social value judgements: principles for the development of NICE guidance – Second edition [Internet]. [cited 2019 May 22]. Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/Research-and-development/Social-Value-Judgements-principles-for-the-development-of-NICE-guidance.pdf
26. Watts S, Stenner P. Doing Q methodology research: Theory, Method and Interpretation. London: Sage Publications Ltd; 2012.
27. Motivaction. The Mentality groups [Internet]. [cited 2018 Feb 22]. Available from: https://www.motivaction.nl/en/mentality/the-mentality-%0Agroups
28. Zorginstituut Nederland. Package advice in practice: Deliberations for arriving at a fair package. Diemen; 2017.
29. Kleinhout-Vliek T, de Bont A, Boer B. The bare necessities? A realist review of necessity argumentations used in health care coverage decisions. Health Policy. 2017;121(7):731–44.
30. Brown S. Political subjectivity: Applications of Q methodology in political science. Connecticut: Yale University Press; 1980.

31. Zabala A. Qmethod: A package to explore human perspectives using Q methodology. R J. 2014;6(2):163–73.
32. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. New Jersey: Lawrence Erlbaum; 1988.
33. Hackert M, Brouwer W, Hoefman R, van Exel N. Views of older people in the Netherlands on wellbeing: A Q-methodology study. Soc Sci Med. 2019;240:112535.

Appendix 7.1: Domains and characteristics of the statements

Table 7.1.1 Statements according to their domains and characteristics

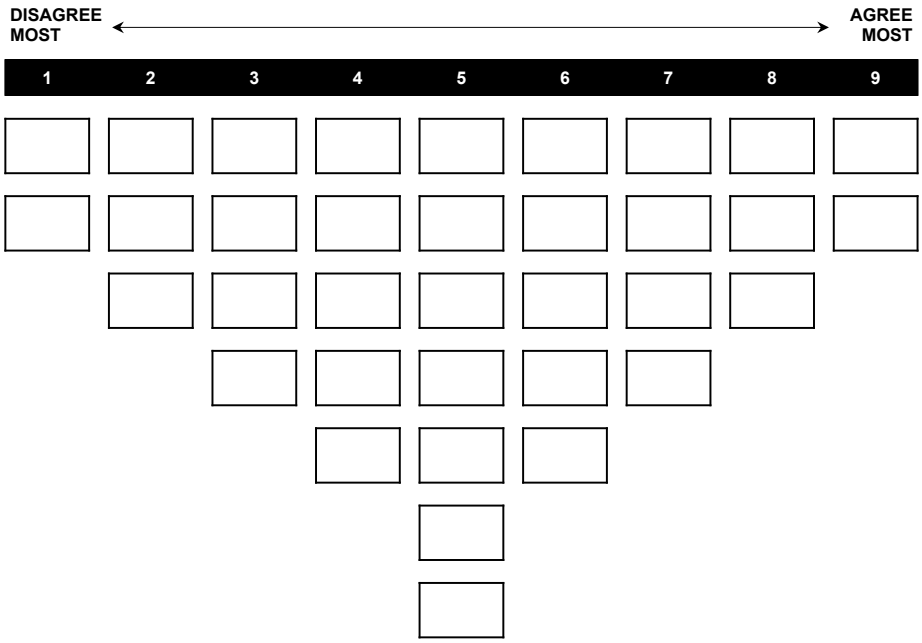
Domain	Characteristic	Statement #
A. Characteristics of the patient	1. Age(ism)/fair innings	1, 18, 2
B. Characteristics of the illness	2. Severity	20
	3. Rarity	33, 35
	4. Rule of rescue	29
	5. Probable cause/culpability	19, 28, 3
C. Characteristics of the treatment	6. Availability	21, 10
	7. Costs/budget impact	30, 31
	8. Efficiency	9
D. Health effects of treatment	9. Size of the effect	5, 27, 4, 23
	10. Length vs. quality of life	22
	11. Start-point before/end-point after treatment	6
	12. Direction of the effect: health gain/loss avoidance	7
	13. Supply induced demand	34
E. Broader effects of treatment	14. Being dependent/caregiving effect	24
	15. Having dependents/family effect/productivity	8, 32
	16. Dignified end of life	26
F. Moral principles	17. Patient choice	11, 12
	18. Values	17, 14, 25
	19. Income/contribution	13
	20. Equality	16, 15

Appendix 7.2: Instruction for participants

These instructions will guide you through the study in a step-by-step manner. Before you start, please read every step in its entirety.

1. The healthcare budget is limited and hence choices have to be made about which treatments can and which cannot be reimbursed from the basic health-insurance package. Shortly, you will read 35 statements about how to allocate the available healthcare resources in an optimal way. We will ask you to what extent you agree with these statements. As this study focuses on your personal view, there will be no right or wrong answers.
2. Please read the statements carefully and divide them into three piles: one pile for statements with which you agree, one pile for statements with which you disagree, and one pile for statements with which you do not explicitly agree or disagree, which are unclear, or which you regard as irrelevant. The numbers on the cards (from 1 to 35) have no other meaning than to help you complete the exercise.
3. Take the pile with statements with which you "AGREE". Read these statements again and select the two statements with which you "AGREE MOST". Place these two statements in the two rightmost boxes on the large sorting grid, below the "9". It does not matter which of the two statements you place at the top or at the bottom. Next, select from the remaining pile of statements with which you "AGREE" the three statements with which you agree most and place them in the three boxes below the "8". Continue with this until all statements with which you "AGREE" have been placed on the sorting grid.
4. Take the pile with statements with which you "DISAGREE". Read these statements again and, like in the previous step, select the two statements with which you "DISAGREE MOST". Place these two statements in the two leftmost boxes on the sorting grid, below the "1". It does not matter which of the two statements you place at the top or at the bottom. Like in the previous step, continue with this until all statements with which you "DISAGREE" have been placed on the sorting grid.
5. Now take the remaining pile with statements. Read these statements again and place them in the remaining open spaces in the middle of the sorting grid, in the way you think they should be placed.
6. You have now placed all the statements on the sorting grid. Read them all again carefully and change the placement of statements if you feel the need to do so.
7. Finally, complete the questions on both pages of the response sheet. Please leave the sorting grid and the statements as they are, we will document this for you.

Appendix 7.3: Sorting grid



Appendix 7.4: (Pearson) correlation matrix for statement rankings at T0 (id#_0) and T1 (id#_1)

Table 7.4.1 (Pearson) correlation matrix for participant's statement rankings at T0 (id#_0) and T1 (id#_1)

id	1_0	2_0	3_0	4_0	5_0	6_0	7_0	8_0	9_0	10_0	11_0	12_0	13_0	14_0	15_0	16_0	17_0	18_0	19_0	20_0	21_0	22_0	23_0	24_0
1_1	0,54	0,38	0,50	0,57	0,70	0,50	0,57	0,26	0,07	0,02	0,05	0,19	0,66	0,41	0,33	0,68	0,55	0,53	0,53	0,41	-0,04	0,69	0,06	0,68
2_1	0,24	0,59	0,25	0,38	0,25	0,12	0,56	0,33	-0,02	0,13	0,06	0,04	0,29	0,23	0,43	0,32	0,11	0,55	0,31	-0,03	0,02	0,15	0,25	0,36
3_1	0,34	0,38	0,51	0,64	0,16	0,45	0,48	0,26	-0,11	0,03	0,48	0,48	0,13	0,18	0,44	0,33	0,16	0,39	0,56	0,17	0,39	0,43	0,40	0,39
4_1	0,28	0,28	0,48	0,65	0,31	0,28	0,55	0,49	-0,03	0,18	0,14	0,26	0,31	0,18	0,43	0,45	0,47	0,54	0,47	0,19	0,10	0,34	0,12	0,44
5_1	0,65	0,76	0,49	0,58	0,73	0,34	0,79	0,50	-0,14	0,04	0,07	0,07	0,65	0,49	0,33	0,69	0,53	0,79	0,53	0,28	-0,01	0,62	0,20	0,56
6_1	0,31	0,25	0,38	0,43	0,17	0,62	0,38	0,27	-0,06	-0,13	0,34	0,47	0,09	0,07	0,38	0,34	0,29	0,33	0,59	0,16	0,39	0,41	0,49	0,28
7_1	0,35	0,19	0,38	0,39	0,13	-0,01	0,59	0,76	-0,32	-0,25	0,24	0,49	0,11	-0,17	0,51	0,44	0,33	0,41	0,49	0,11	0,51	0,20	0,29	0,26
8_1	0,07	0,19	0,44	0,50	0,17	0,37	0,31	0,19	-0,31	-0,06	0,29	0,53	0,13	-0,01	0,44	0,30	0,37	0,32	0,57	0,23	0,21	0,23	0,26	0,33
9_1	0,13	-0,10	0,19	0,04	-0,24	0,09	0,00	0,13	0,26	0,04	0,01	0,16	-0,18	-0,19	0,24	0,09	-0,01	-0,04	0,06	-0,25	0,33	-0,14	0,16	-0,21
10_1	-0,07	0,30	-0,19	0,14	0,07	0,11	0,06	-0,26	0,36	0,78	-0,30	-0,21	0,18	0,51	-0,14	0,03	0,11	0,11	0,11	-0,04	-0,19	-0,11	-0,06	0,14
11_1	0,34	0,17	0,42	0,46	0,06	0,36	0,34	0,51	0,10	-0,01	0,41	0,47	0,07	-0,06	0,49	0,31	0,11	0,29	0,54	0,23	0,54	0,15	0,59	0,20
12_1	0,00	-0,22	-0,01	0,01	-0,46	0,09	0,05	0,46	-0,20	-0,11	0,41	0,43	-0,33	-0,48	0,26	-0,21	-0,21	-0,07	0,26	-0,01	0,63	-0,21	0,39	-0,11
13_1	0,46	0,53	0,27	0,50	0,68	0,56	0,59	-0,04	0,06	0,14	-0,07	-0,04	0,67	0,49	0,18	0,66	0,54	0,63	0,39	0,20	-0,22	0,54	0,11	0,69
14_1	0,37	0,64	0,09	0,39	0,49	0,36	0,48	0,01	0,17	0,46	-0,21	-0,21	0,48	0,60	-0,11	0,44	0,49	0,59	0,32	-0,02	-0,32	0,36	-0,01	0,45
15_1	0,32	0,34	0,55	0,54	0,33	0,28	0,68	0,49	-0,22	-0,06	0,34	0,39	0,44	0,13	0,77	0,56	0,30	0,53	0,65	0,45	0,28	0,36	0,49	0,64
16_1	0,63	0,30	0,38	0,41	0,44	0,28	0,60	0,56	-0,29	-0,16	0,06	0,27	0,24	0,16	0,28	0,56	0,44	0,57	0,43	0,13	0,23	0,50	0,39	0,33
17_1	0,01	-0,09	0,16	-0,07	0,04	0,05	0,08	0,08	-0,19	-0,20	-0,14	0,28	-0,02	-0,13	0,36	0,22	0,21	0,11	0,21	0,15	0,07	0,07	0,18	0,28
18_1	0,40	0,49	0,31	0,58	0,61	0,24	0,76	0,43	-0,06	0,23	0,00	-0,04	0,65	0,51	0,24	0,56	0,49	0,78	0,44	0,17	-0,16	0,59	0,04	0,64
19_1	0,38	0,31	0,46	0,42	0,08	0,21	0,53	0,42	-0,29	-0,11	0,46	0,61	0,04	0,01	0,63	0,30	0,11	0,34	0,71	0,31	0,64	0,23	0,69	0,32
20_1	0,20	0,19	0,34	0,26	0,01	0,28	0,25	0,33	-0,39	-0,07	0,54	0,63	0,14	-0,08	0,54	0,13	0,20	0,22	0,59	0,62	0,46	0,12	0,63	0,39
21_1	0,09	-0,07	0,27	0,14	-0,33	0,09	0,09	0,20	-0,22	-0,23	0,35	0,64	-0,34	-0,29	0,44	-0,02	-0,06	-0,08	0,46	0,04	0,76	-0,14	0,47	-0,13
22_1	0,43	0,43	0,49	0,56	0,61	0,42	0,59	0,34	-0,16	-0,01	0,09	0,05	0,57	0,29	0,39	0,60	0,49	0,69	0,49	0,33	-0,03	0,64	0,18	0,71
23_1	0,14	-0,20	0,46	0,27	-0,14	0,14	0,16	0,33	-0,11	-0,19	0,64	0,76	-0,01	-0,28	0,64	0,06	0,03	0,01	0,44	0,44	0,63	0,02	0,49	0,20
24_1	0,51	0,45	0,30	0,38	0,42	0,13	0,73	0,44	-0,28	-0,19	0,11	0,03	0,38	0,15	0,32	0,49	0,31	0,69	0,37	0,13	-0,01	0,46	0,19	0,63

Appendix 7.5: (Pearson) correlation matrix for statement rankings at T0 (id#_0)

Table 7.5.1 (Pearson) correlation matrix for participant's statement rankings at T0 (id#_0)

id	1_0	2_0	3_0	4_0	5_0	6_0	7_0	8_0	9_0	10_0	11_0	12_0	13_0	14_0	15_0	16_0	17_0	18_0	19_0	20_0	21_0	22_0	23_0	24_0
1_0	1.00	0.44	0.48	0.46	0.60	0.22	0.58	0.36	-0.02	-0.17	0.09	0.11	0.53	0.39	0.27	0.73	0.50	0.48	0.31	0.26	0.16	0.59	0.25	0.43
2_0	0.44	1.00	0.39	0.54	0.66	0.27	0.56	0.08	-0.01	0.23	-0.08	-0.05	0.52	0.65	0.28	0.59	0.39	0.67	0.50	0.19	-0.16	0.48	0.13	0.48
3_0	0.48	0.39	1.00	0.67	0.46	0.38	0.54	0.33	-0.03	-0.21	0.54	0.44	0.34	0.08	0.61	0.59	0.29	0.52	0.60	0.35	0.28	0.51	0.39	0.47
4_0	0.46	0.54	0.67	1.00	0.50	0.41	0.61	0.29	0.09	0.19	0.39	0.31	0.52	0.38	0.51	0.57	0.41	0.59	0.63	0.30	0.19	0.55	0.16	0.61
5_0	0.60	0.66	0.46	0.50	1.00	0.36	0.58	0.13	-0.05	0.08	-0.14	-0.11	0.81	0.59	0.11	0.79	0.64	0.71	0.40	0.36	-0.36	0.68	0.07	0.66
6_0	0.22	0.27	0.38	0.41	0.36	1.00	0.26	-0.07	0.11	0.07	0.13	0.20	0.29	0.16	0.17	0.39	0.37	0.39	0.49	0.22	-0.01	0.50	0.19	0.43
7_0	0.58	0.56	0.54	0.61	0.58	0.26	1.00	0.49	-0.06	-0.04	0.21	0.18	0.59	0.34	0.54	0.77	0.49	0.76	0.58	0.16	0.14	0.52	0.30	0.65
8_0	0.36	0.08	0.33	0.29	0.13	-0.07	0.49	1.00	-0.31	-0.20	0.30	0.34	0.17	-0.17	0.36	0.28	0.16	0.40	0.37	0.17	0.42	0.23	0.30	0.18
9_0	-0.02	-0.01	-0.03	0.09	-0.05	0.11	-0.06	-0.31	1.00	0.41	-0.04	-0.33	0.12	0.22	0.04	-0.01	-0.01	-0.05	-0.22	-0.24	-0.23	-0.16	-0.09	-0.04
10_0	-0.17	0.23	-0.21	0.19	0.08	0.07	-0.04	-0.20	0.41	1.00	-0.16	-0.28	0.26	0.46	-0.21	-0.13	0.18	0.14	0.03	0.03	-0.29	-0.11	-0.05	0.09
11_0	0.09	-0.08	0.54	0.39	-0.14	0.13	0.21	0.30	-0.04	-0.16	1.00	0.63	0.04	-0.28	0.58	-0.08	-0.14	0.07	0.41	0.44	0.60	0.09	0.49	0.16
12_0	0.11	-0.05	0.44	0.31	-0.11	0.20	0.18	0.34	-0.33	-0.28	0.63	1.00	-0.10	-0.31	0.56	0.09	0.04	-0.01	0.52	0.43	0.71	0.08	0.40	0.13
13_0	0.53	0.52	0.34	0.52	0.81	0.29	0.59	0.17	0.12	0.26	0.04	-0.10	1.00	0.60	0.19	0.65	0.63	0.62	0.34	0.49	-0.26	0.61	0.10	0.71
14_0	0.39	0.65	0.08	0.38	0.59	0.16	0.34	-0.17	0.22	0.46	-0.28	-0.31	0.60	1.00	-0.05	0.45	0.49	0.39	0.17	0.08	-0.43	0.44	-0.13	0.28
15_0	0.27	0.28	0.61	0.51	0.11	0.17	0.54	0.36	0.04	-0.21	0.58	0.56	0.19	-0.05	1.00	0.41	0.05	0.34	0.58	0.33	0.52	0.20	0.55	0.47
16_0	0.73	0.59	0.59	0.57	0.79	0.39	0.77	0.28	-0.01	-0.13	-0.08	0.09	0.65	0.45	0.41	1.00	0.64	0.67	0.49	0.21	-0.03	0.59	0.20	0.68
17_0	0.50	0.39	0.29	0.41	0.64	0.37	0.49	0.16	-0.01	0.18	-0.14	0.04	0.63	0.49	0.05	0.64	1.00	0.56	0.40	0.24	-0.26	0.46	0.07	0.51
18_0	0.48	0.67	0.52	0.59	0.71	0.39	0.76	0.40	-0.05	0.14	0.07	-0.01	0.62	0.39	0.34	0.67	0.56	1.00	0.60	0.17	-0.18	0.59	0.23	0.71
19_0	0.31	0.50	0.60	0.63	0.40	0.49	0.58	0.37	-0.22	0.03	0.41	0.52	0.34	0.17	0.58	0.49	0.40	0.60	1.00	0.46	0.43	0.42	0.51	0.55
20_0	0.26	0.19	0.35	0.30	0.36	0.22	0.16	0.17	-0.24	0.03	0.44	0.43	0.49	0.08	0.33	0.21	0.24	0.17	0.46	1.00	0.20	0.26	0.34	0.50
21_0	0.16	-0.16	0.28	0.19	-0.36	-0.01	0.14	0.42	-0.23	-0.29	0.60	0.71	-0.26	-0.43	0.52	-0.03	-0.26	-0.18	0.43	0.20	1.00	-0.07	0.56	-0.11
22_0	0.59	0.48	0.51	0.55	0.68	0.50	0.52	0.23	-0.16	-0.11	0.09	0.08	0.61	0.44	0.20	0.59	0.46	0.59	0.42	0.26	-0.07	1.00	0.11	0.53
23_0	0.25	0.13	0.39	0.16	0.07	0.19	0.30	0.30	-0.09	-0.05	0.49	0.40	0.10	-0.13	0.55	0.20	0.07	0.23	0.51	0.34	0.56	0.11	1.00	0.19
24_0	0.43	0.48	0.47	0.61	0.66	0.43	0.65	0.18	-0.04	0.09	0.16	0.13	0.71	0.28	0.47	0.68	0.51	0.71	0.55	0.50	-0.11	0.53	0.19	1.00

Appendix 7.6: (Pearson) correlation matrix for statement rankings at T1 (id#_1)

Table 7.6.1 (Pearson) correlation matrix for participant's statement rankings at T_1 (id#_1)

id	1_1	2_1	3_1	4_1	5_1	6_1	7_1	8_1	9_1	10_1	11_1	12_1	13_1	14_1	15_1	16_1	17_1	18_1	19_1	20_1	21_1	22_1	23_1	24_1
1_1	1.00	0.14	0.25	0.53	0.58	0.30	0.23	0.21	0.01	0.19	0.25	-0.26	0.51	0.26	0.44	0.44	0.25	0.65	0.28	0.08	-0.01	0.62	0.24	0.34
2_1	0.14	1.00	0.32	0.41	0.59	0.26	0.30	0.21	0.12	0.21	0.36	0.24	0.37	0.43	0.50	0.38	0.07	0.50	0.49	0.27	0.21	0.29	0.03	0.54
3_1	0.25	0.32	1.00	0.43	0.38	0.60	0.34	0.39	0.00	0.04	0.48	0.33	0.44	0.46	0.47	0.48	-0.07	0.36	0.66	0.53	0.30	0.36	0.41	0.47
4_1	0.53	0.41	0.43	1.00	0.45	0.19	0.59	0.41	0.34	0.12	0.46	0.10	0.35	0.26	0.63	0.65	0.30	0.67	0.44	0.21	0.23	0.61	0.39	0.37
5_1	0.58	0.59	0.38	0.45	1.00	0.36	0.46	0.24	-0.11	0.06	0.30	-0.03	0.50	0.52	0.50	0.58	-0.06	0.73	0.39	0.28	-0.04	0.61	-0.06	0.68
6_1	0.30	0.26	0.60	0.19	0.36	1.00	0.36	0.44	-0.06	0.00	0.61	0.41	0.33	0.25	0.36	0.43	0.16	0.18	0.55	0.48	0.36	0.24	0.27	0.21
7_1	0.23	0.30	0.34	0.59	0.46	0.36	1.00	0.39	0.29	-0.23	0.51	0.48	0.11	0.06	0.56	0.62	0.34	0.40	0.50	0.38	0.36	0.39	0.35	0.42
8_1	0.21	0.21	0.39	0.41	0.24	0.44	0.39	1.00	-0.03	-0.17	0.31	0.19	0.22	0.09	0.51	0.23	0.18	0.11	0.42	0.49	0.42	0.37	0.40	0.21
9_1	0.01	0.12	0.00	0.34	-0.11	-0.06	0.29	-0.03	1.00	0.12	0.28	0.29	-0.11	-0.14	0.13	0.16	0.26	-0.08	0.13	-0.03	0.43	-0.07	0.27	-0.21
10_1	0.19	0.21	0.04	0.12	0.06	0.00	-0.23	-0.17	0.12	1.00	0.04	-0.13	0.22	0.43	-0.01	-0.07	0.02	0.25	0.07	-0.08	-0.03	-0.06	-0.16	-0.18
11_1	0.25	0.36	0.48	0.46	0.30	0.61	0.51	0.31	0.28	0.04	1.00	0.54	0.13	0.21	0.44	0.59	0.13	0.25	0.66	0.44	0.51	0.24	0.46	0.20
12_1	-0.26	0.24	0.33	0.10	-0.03	0.41	0.48	0.19	0.29	-0.13	0.54	1.00	-0.20	-0.03	0.24	0.24	0.11	-0.08	0.49	0.52	0.49	-0.16	0.33	0.16
13_1	0.51	0.37	0.44	0.35	0.50	0.33	0.11	0.22	-0.11	0.22	0.13	-0.20	1.00	0.64	0.38	0.44	0.01	0.54	0.19	0.14	-0.14	0.64	0.06	0.41
14_1	0.26	0.43	0.46	0.26	0.52	0.25	0.06	0.09	-0.14	0.43	0.21	-0.03	0.64	1.00	0.18	0.33	-0.21	0.51	0.23	0.09	-0.21	0.29	-0.26	0.49
15_1	0.44	0.50	0.47	0.63	0.50	0.36	0.56	0.51	0.13	-0.01	0.44	0.24	0.38	0.18	1.00	0.47	0.40	0.53	0.64	0.64	0.22	0.53	0.47	0.51
16_1	0.44	0.38	0.48	0.65	0.58	0.43	0.62	0.23	0.16	-0.07	0.59	0.24	0.44	0.33	0.47	1.00	0.24	0.60	0.62	0.28	0.29	0.53	0.27	0.55
17_1	0.25	0.07	-0.07	0.30	-0.06	0.16	0.34	0.18	0.26	0.02	0.13	0.11	0.01	-0.21	0.40	0.24	1.00	0.13	0.26	0.28	0.27	0.20	0.36	0.02
18_1	0.65	0.50	0.36	0.67	0.73	0.18	0.40	0.11	-0.08	0.25	0.25	-0.08	0.54	0.51	0.53	0.60	0.13	1.00	0.33	0.09	-0.13	0.63	-0.02	0.66
19_1	0.28	0.49	0.66	0.44	0.39	0.55	0.50	0.42	0.13	0.07	0.66	0.49	0.19	0.23	0.64	0.62	0.26	0.33	1.00	0.66	0.69	0.28	0.59	0.48
20_1	0.08	0.27	0.53	0.21	0.28	0.48	0.38	0.49	-0.03	-0.08	0.44	0.52	0.14	0.09	0.64	0.28	0.28	0.09	0.66	1.00	0.41	0.24	0.56	0.33
21_1	-0.01	0.21	0.30	0.23	-0.04	0.36	0.36	0.42	0.43	-0.03	0.51	0.49	-0.14	-0.21	0.22	0.29	0.27	-0.13	0.69	0.41	1.00	0.02	0.65	-0.01
22_1	0.62	0.29	0.36	0.61	0.61	0.24	0.39	0.37	-0.07	-0.06	0.24	-0.16	0.64	0.29	0.53	0.53	0.20	0.63	0.28	0.24	0.02	1.00	0.26	0.49
23_1	0.24	0.03	0.41	0.39	-0.06	0.27	0.35	0.40	0.27	-0.16	0.46	0.33	0.06	-0.26	0.47	0.27	0.36	-0.02	0.59	0.56	0.65	0.26	1.00	-0.01
24_1	0.34	0.54	0.47	0.37	0.68	0.21	0.42	0.21	-0.21	-0.18	0.20	0.16	0.41	0.49	0.51	0.55	0.02	0.66	0.48	0.33	-0.01	0.49	-0.01	1.00

7

Chapter 8
Discussion

8.1 Introduction

The increasing demand for healthcare and the resulting pressure on limited budgets renders priority setting in the allocation of healthcare resources inevitable. It has been argued that an optimal allocation of healthcare resources involves setting priorities that contribute to meeting efficiency as well as equity objectives of publicly financed healthcare systems. Whilst an increasing number of countries integrates societal concerns for equity with concerns for efficiency into the decision-making framework, important questions remain about what and how equity considerations should be integrated and what (relative) weight these considerations should receive in resource-allocation decisions. This thesis addressed these questions with the objective to contribute to the improvement of the decision-making framework by providing further insight into societal concerns for equity, in particular for priority setting based on disease severity and the age of patients.

This final chapter discusses the main findings of this thesis in relation to the research questions outlined in Chapter 1. Furthermore, this chapter discusses the strengths and limitations of this thesis and highlights several of its implications for policy and future research.

8.2 Main findings

Chapter 2 examined the normative justification and empirical support for equity weighting based on proportional shortfall in the Netherlands (research question 1). The results of this chapter indicated that the decision model in which severity-based equity weights are applied in economic evaluations is generally supported and increasingly used for healthcare priority setting in the Netherlands. However, the results also indicated that theoretical and empirical support for defining disease severity in terms of proportional shortfall may be limited. Proportional shortfall combines aspects of the renowned severity and fair innings approaches and may, therefore, insufficiently reflect societal preferences for either of these two approaches in resource-allocation decisions. Proportional shortfall may further insufficiently reflect age-related preferences in society, in particular for prioritisation of younger patients in resource-allocation decisions. Empirical evidence on the relationship between societal preferences for priority setting based on disease severity (e.g. defined in terms of proportional shortfall) and the age of patients is scarce. Therefore, we conducted a series of empirical studies to obtain further insight into the relative strength of these preferences. Chapters 3 to 5 described the results of these studies.

Chapter 3 examined how much weight members of the public the Netherlands attach to disease severity and the age of patients in resource-allocation decisions (research question 2). We applied the person trade-off approach in an innovative manner to estimate separate *and* combined weights for priority

setting based on disease severity and age. When estimating these weights separately, we found that respondents, on average, attached a higher weight to allocating resources towards more severely ill patients and towards younger patients. When estimating these weights jointly, we found that respondents, on average, attached a higher weight to allocating towards younger patients, even when they were less severely ill than older patients. We further found that the implied equity weights were dependent on the type of difference (i.e. in terms of disease severity and/or age) and size of the difference between the patient groups. Chapters 4 and 5 examined what members of the public the Netherlands are willing to pay for health gains in patients with different ages and levels of disease severity (research question 3), the latter also operationalised in an end-of-life context (research question 4). In accordance with the results of the previous chapter, we found that respondents, on average, had a higher willingness to pay for health gains in younger patients and in more severely ill patients, though at a decreasing marginal rate. We found that the willingness to pay for health gains in patients of different ages was dependent on their disease severity and the size of the health gain. We further found that the willingness to pay was higher for health gains in patients who fully recovered than for similar health gains in patients who died one year after falling ill. In addition, we found some evidence that willingness to pay for health gains in end-of-life patients was dependent on the type of health gain as willingness to pay was, on average, higher for gains in health-related quality of life than for gains in life expectancy. A consistent finding in Chapters 3 to 5 was that respondents, on average, least preferred to allocate resources towards health gains in the oldest patient group (i.e. patients aged 70). Another consistent finding was that there is considerable preference heterogeneity amongst members of the public, both in the strength and (to a lesser extent) direction of their preferences.

Chapter 6 examined how different viewpoints on healthcare priority setting relate to concerns for equity and efficiency in resource-allocation decisions (research question 5). We found that the vast majority of respondents believed that access to healthcare is a basic human right to which patients are entitled without restriction. Respondents who held this view opposed priority setting based on characteristics of the patients, disease, and health technology and believed that priority setting should solely be based on patients' need for care. We found that a minority of respondents appreciated that healthcare resources are scarce and priority setting is necessary. Of those who held this latter view, most believed that priority setting should be based on the type and size of health gains that health technologies generate and that resources should primarily be allocated towards health technologies that increase patients' quality of life. Only a few respondents believed that priority setting should be based on patient characteristics and the cost-effectiveness of health technologies. The findings of Chapter 6 further indicated that members of the public who oppose priority setting based on anything other than patients' need for care are less likely to prioritise between competing

health technologies and patient populations than others. However, this was dependent of the specific context of the resource-allocation decision. Indeed, most respondents were willing to prioritise one patient group over another at least once, irrespective of their view on healthcare priority setting.

Finally, in Chapter 7 we applied Q methodology in an innovative manner to examine whether and how participating in a deliberative citizens panel influenced participants' viewpoints on healthcare priority setting (research question 6). The results of this chapter indicated that participants' initial viewpoints partly remained stable over the course of the panel. In particular, participants' belief that access to healthcare is a basic human right and that priority setting should be based on patients' need for care and the type and size of health gains remained important. Notable changes in participants' viewpoints concerned their increased support for healthcare priority setting, consideration of healthcare costs, and appreciation of cost-effectiveness as a relevant decision criterion after participating in the panel. Despite these changes, we found that the viewpoints of participants only moderately converged and remained diverse after the panel.

The findings of this thesis provide further insight into societal concerns for equity, in particular for priority setting based on disease severity and the age of patients. The findings indicate that applying equity weights based on proportional shortfall in economic evaluations is largely consistent with societal preferences for prioritising more severely ill patients; however, may be insufficiently consistent with societal preferences for prioritisation of younger patients in resource-allocation decisions. The insights provided by this thesis, for example into the rather complex relationship between societal preferences for prioritising more severely ill and younger patients, can be used to inform decisions on improving (different key aspects of) the current decision-making framework in the Netherlands. Furthermore, the insights can be used to inform decisions on how to involve the public in the decision-making process, be it on the level of communication, consultation, or participation.

8.3 Strengths and limitations

The main strength of this thesis lies in the application of different methods for obtaining insight into societal concerns for equity in healthcare priority setting. Firstly, we applied different methods in order to elicit different (i.e. non-monetary and monetary) types of societal preferences. Hence, the findings of this thesis can be used to improve different key aspects of the decision-making framework in the Netherlands. Secondly, we applied different methods—whilst standardising certain design aspects to improve comparability between findings—in order to examine the robustness of societal concerns for equity across preference-elicitation and statistical methods. Although each chapter of this thesis provides some new insights into societal concerns for equity, the chapters also confirm the main findings that in general, members of the public

in the Netherlands, seem to prefer allocating healthcare resources towards more severely ill *and* younger patients. Therefore, the empirical evidence presented in this thesis may be considered compelling. Finally, we applied the person trade-off method and Q methodology in innovative manners in order to examine to what extent societal preferences for priority setting relate to disease severity and the age of patients (Chapter 3) and how viewpoints on healthcare priority setting are influenced by participating in a deliberative citizens panel on this topic (Chapter 7). As such, this thesis contributes to the further development and wider application of these two methods in, and possibly beyond, the context of healthcare priority setting. Despite this strength, some general limitations need to be taken into account when interpreting the findings and considering the policy and research implications of this thesis.

A first limitation concerns an important premise that underlies the studies presented in this thesis, specifically that the decision-making framework should (to some extent or entirely) be aligned with societal preferences for healthcare priority setting. It should be noted that whether and, if so, to what extent resource-allocation decisions in publicly financed healthcare systems should be aligned with the preferences of members of the general public are in themselves normative questions that cannot be answered with empirical research. Normative research is warranted to examine if this alignment is indeed desirable or if resource-allocation decisions should, for example, (also) be aligned with the preferences of other stakeholders, such as patients and decision makers. Irrespective of the outcome of such research, the importance of understanding and (in some way) accounting for societal preferences in resource-allocation decisions may be evident from the public debate and controversy that often follow (in particular, negative) decisions. This may further be evident from the increased interest amongst decision makers in the participation of members of the public in resource-allocation decisions. The empirical studies presented in this thesis may, therefore, aid decision makers in at least two ways. Firstly, the findings may aid them in better aligning the outcomes and process of resource-allocation decisions with societal preferences. Secondly, the findings may aid them in identifying potential controversies and communicating decisions on the allocation of resources or on adjustments to the decision-making framework to the public.

A second limitation concerns the focus of this thesis on the examination of societal preferences for priority setting based on, specifically, disease severity and the age of patients (Chapters 2 to 5). This focus was intentional and enabled us to provide detailed insight into the complex relationship between these societal preferences. Nonetheless, this focus inevitably resulted in largely disregarding other (potentially relevant) societal preferences for priority setting. For example, societal preferences for priority setting based on the prevalence of a disease (i.e. whether it is common or rare) or based on the socioeconomic status, productivity, individual responsibility (e.g. illness due to unhealthy lifestyle choices), or social role of patients have not been exa-

mined in this thesis. Further research is warranted to examine (the strength of) such other societal preferences for priority setting, also in relation to disease severity and the age of patients.

A third limitation, related to the narrow focus discussed above, concerns the elicitation of societal preferences for priority setting based on disease severity and the age of patients under the assumption that the health gains and costs of health technologies were the same for all patients (Chapters 3 to 5). Ideally, we would have examined these preferences in relation to differences in costs and benefits, and hence to differences in the level of cost-effectiveness of health technologies. This would have resulted in more realistic scenarios, albeit potentially to the detriment of our ability to disentangle respondents' reasons for preferring one health technology or patient group over another. Further research is warranted to examine societal preferences for priority setting, based on not only disease severity and the age of patients, but also based on other characteristics of the patients, diseases, and health technologies, including their level of cost-effectiveness.

A fourth limitation concerns the collection of data at single points in time for obtaining insight into societal preferences for priority setting (Chapters 3 to 6). The applied designs facilitated a thorough examination of societal preferences based on cross-sectional as well as panel data structures. However, the designs did not facilitate the examination of consistencies or of possible changes in societal preferences over time. Just like viewpoints on healthcare priority setting may change over time, for example, under the influence of participating in a citizens panel on this topic (Chapter 7), societal preferences for priority setting may also change. The collection of data at multiple points in time may aid in obtaining insight into potential changes in societal preferences as well as into the underlying causal mechanisms. Under the assumption that decision makers (at least to some extent) aim to align the outcomes and process of resource-allocation decisions with societal preferences, regular assessment of societal preferences may help decision makers in communicating decisions to the public and in adjusting the decision-making framework in accordance with changing preferences. Note that such adjustments need to be weighed against preferences for consistency in decision making over time.

A final limitation concerns the external validity of the empirical studies presented in this thesis. Firstly, we elicited societal preferences in samples drawn from (online) research panels. As such, the generalisability of our findings need to be considered in light of the inevitable risk of selection and non-response bias that comes with the use of such panels. Secondly, we elicited societal preferences in quota samples that were representative of the general public in terms of age (i.e. 18+), sex, and education level (in Chapters 3 to 6) and, in addition, of the diversity in values, lifestyles, and consumption patterns (in Chapter 7) in the Netherlands. Evidently, the generalisation of our findings to populations outside of the Netherlands warrants caution, for

example, because of socioeconomic and cultural differences between countries. Multi-country data may be used to obtain insight into the impact of such differences and the generalisability of our findings to resource-allocation decisions in other countries as well as to decisions taken at a supranational level. Note that members of the public in the Netherlands may, on average, be more accustomed to more explicit public debates about sensitive and controversial issues, such as healthcare priority setting, than in some other countries. Consequently, they may also be more willing to and comfortable with completing preference-elicitation tasks in which their preferences are elicited in a fairly direct manner (e.g. person trade-off and contingent-valuation tasks). Researchers should be aware that the elicitation of societal preferences in other cultural contexts may require different methods than those applied in this thesis. Finally, we elicited societal preferences under certainty in relatively simple hypothetical resource-allocation scenarios. This may have increased the internal validity of our research, for example, by improving the feasibility of the preference-elicitation tasks for respondents. However, this inevitably came to the detriment of the generalisability of our findings to the complex resource-allocation decisions with which decision makers are typically confronted in real life.

8.4 Implications for policy and future research

This thesis provided further insight into societal concerns for equity with the objective to improve priority setting in healthcare. The findings of this thesis have several implications for policy and future research.

Several countries have incorporated societal preferences for priority setting based on disease severity into the decision-making framework. In the Netherlands, disease severity is defined in terms of proportional shortfall and divided into four severity classes with the successive proportional-shortfall levels of less than 10%, between 10 and 40%, between 41 and 70% and more than 71%. Resources are, in principle, not allocated towards health technologies indicated for patients who fall into the lowest class (i.e. suggesting a monetary threshold of €0 in economic evaluations). The incremental cost-effectiveness ratios of health technologies indicated for patients who fall into the three higher classes are evaluated against monetary thresholds of €20,000, €50,000, and €80,000, respectively. The findings of this thesis indicate that the principles underlying this framework may not be fully aligned with societal preferences. For example, they indicate that the public may indeed attach a higher weight to health gains in more severely ill patients; however, the differences between the monetary thresholds applied for the different severity classes may be smaller (at least in relative terms) than in the current decision-making framework. It needs emphasis that the empirical studies presented in this thesis examined societal preferences in such relative terms, and hence their findings cannot fully inform decisions on the absolute height of

the monetary threshold. Further empirical research is warranted to examine appropriate monetary thresholds in absolute terms. Further methodological research is also warranted to examine what approach would be most suitable for examining this.

In addition to societal preferences for priority setting based on disease severity, decision makers may consider incorporating societal preferences for priority setting based on the age of patients into the decision-making framework for two main reasons. Firstly, accumulating evidence shows that the public considers age to be an important equity-relevant characteristic, perhaps even more important than disease severity. It could be argued that these societal preferences should be taken into account in decisions that concern the allocation of public healthcare resources, at least in communicating (the reasons that underlie) such decisions to the public. Secondly, definitions of disease severity that are derived from the severity and fair innings approaches are not independent from age. Indeed, some age groups may implicitly be prioritised over other age groups, even when decision makers do not (wish to) explicitly account for age in resource-allocation decisions. As such, the paradox in the Netherlands is that proportional shortfall may need to be corrected for societal preferences for prioritisation of younger patients, even when the ultimate objective is to avoid discrimination on the basis of age between patients.

The findings of this thesis clearly indicate that the public has a preference for allocating healthcare resources towards younger patients. Considering that we elicited societal preferences in relative terms, we want to stress that the findings should be interpreted with care. Decision makers should, for example, be aware that an age preference as detected in this thesis should not be equated with an absolute preference for allocating healthcare resources towards children. Moreover, the findings also do not allow the conclusion that the public has no preference for allocating resources towards older (e.g. 70-year-old) patients. Decision makers should also be aware that the empirical studies presented in this thesis reveal considerable preference uncertainty and heterogeneity. Hence, aligning the outcomes of resource-allocation decisions with societal preferences will likely remain challenging. Indeed, even when decision makers account for the relatively strong societal preferences for prioritisation of younger patients in resource-allocation decisions, these decisions will likely still be met with opposition from some groups in society and be followed by public debate.

Aligning the outcomes of resource-allocation decisions with societal preferences will likely also remain challenging in consideration of the diversity of societal concerns for equity and their interdependency and dependency on contextual and personal factors. Further research is warranted to examine how the eclectic mix of societal concerns for equity can best be combined and integrated into the decision-making framework, whilst retaining its ability to facilitate transparent and consistent decision making. Irrespective of the

outcome of this research, decision makers may also consider (better) aligning the process of resource-allocation decisions with societal preferences in order to increase the legitimacy of and public support for resource-allocation decisions. Further research is warranted to examine on what level members of the public would want and would be able to be involved in the decision-making process, whether and how this can be operationalised, and what the (e.g. political and societal) impact of this involvement will be.

8.5 Final remarks

The increasing demand for healthcare and the resulting pressure on limited budgets renders healthcare priority setting inevitable. In order to allocate the available resources in an optimal way for society, countries aim to integrate societal concerns for equity and efficiency into the decision-making framework. This thesis contributed to meeting this aim by providing further insight into societal concerns for equity in this context. It showed that members of the public consider priority setting based on disease severity and on the age of patients important in resource-allocation decisions in healthcare. The findings of this thesis suggest that the current decision-making framework (in the Netherlands as well as in other countries) could be adjusted to better reflect age-related preferences in society, not only because these preferences are relatively strong but also because the age of patients is inextricably related to definitions of disease severity. The findings further suggest that, in addition to adjusting the decision-making framework to (better) align the outcomes of decisions with societal preferences, the framework can (also) be adjusted to align the process of decision making with societal preferences. For example, by implementing citizens panels that inform or take on a more prominent role in resource-allocation decisions.

The findings of this thesis moreover suggest that societal concerns for equity are diverse, interdependent, and dependent on contextual and personal factors. Therefore, (better) aligning the outcomes and process of resource-allocation decisions, and hence increasing the legitimacy and public support for these decisions, will likely continue to be a challenge for decision makers. My hope is that this thesis contributes to facing this challenge and ultimately to allocating the available healthcare resources in a more efficient and equitable manner.

8

Summary

The demand for healthcare is rapidly increasing for reasons that include ageing populations, the availability of increasingly advanced and expensive new health technologies, and higher standards of living that raise the expectations of health and healthcare. As healthcare resources are limited, the resulting pressure on available budgets renders priority setting in the allocation of healthcare resources inevitable.

Two important objectives of publicly financed healthcare systems are to generate as much (health) value as possible from the healthcare budget and to distribute health and healthcare fairly. To set priorities that contribute to meeting both objectives, an increasing number of countries—among which the Netherlands—integrates efficiency and equity considerations into their decision-making frameworks. For example, by applying equity weights in economic evaluations of new health technologies. Important questions remain, however, about which equity considerations should be considered, what (relative) weight these considerations should receive, and how they should be incorporated in resource-allocation decisions. This thesis addressed these questions with the objective to contribute to the improvement of decision-making frameworks by providing further insight into societal concerns for equity, in particular for priority setting based on disease severity and the age of patients.

Chapter 2 examined the normative justification and empirical support for equity weighting based on proportional shortfall in the Netherlands. Proportional shortfall is derived from the renowned severity and fair innings approaches and used to define and quantify disease severity in resource-allocation decisions. The results of this chapter indicated that the decision model in which severity-based equity weights are applied in economic evaluations is generally supported and increasingly used for healthcare priority setting in the Netherlands. However, the results also indicated that theoretical and empirical support for defining and quantifying disease severity in terms of proportional shortfall is limited. Proportional shortfall, for example, insufficiently reflected age-related preferences in society, in particular for prioritisation of younger patients in resource-allocation decisions.

Chapters 3 to 5 built on the results of Chapter 2 and examined societal preferences for healthcare priority setting based on disease severity and the age of patients. Chapter 3 presents the results of an experiment in which we estimated separate and combined weights for priority setting based on disease severity and age. The results indicated that respondents, on average, attached a higher weight to allocating resources towards more severely ill and towards younger patients. The results also indicated that preferences for prioritisation of younger patients were relatively strong as respondents attached a higher weight to younger patients, even when they were less severely ill than older

patients. Chapters 4 and 5 present the results of an experiment in which we examined the willingness to pay of respondents for health gains in patients with different ages and levels of disease severity, the latter also operationalised in an end-of-life context. The results of these two chapters indicated that willingness to pay was, on average, higher for health gains in younger and in more severely ill patients, though at a decreasing marginal rate. The results furthermore indicated that willingness to pay for health gains in patients of different ages was dependent on their level of disease severity, the size (and to a lesser extent the type) of the health gain, and the outcome of the disease. We found that willingness to pay was, on average, lower for gains in end-of-life patients than for similar gains in patients who fully recovered.

Chapter 6 examined how different viewpoints on healthcare priority setting related to societal concerns for equity and efficiency in resource-allocation decisions. The results of this chapter indicated that the majority of respondents held a view that opposed priority setting based on characteristics of the patients, disease, and health technology and believed that priority setting should solely be based on patients' need for care. A minority of respondents appreciated the scarcity of healthcare resources and necessity of healthcare priority setting. However, only a few respondents believed that priority setting should be based on patient characteristics and the cost-effectiveness of health technologies. The results indicated that respondents who opposed priority setting were less likely to prioritise between competing health technologies and patient populations than respondents with another view. Nonetheless, the results showed that this was dependent on the specific decision-making context. Indeed, most respondents were willing to prioritise one technology or patient group over another at least once, irrespective of their view on healthcare priority setting.

Chapter 7 examined how participating in a deliberative citizens panel influenced participants' viewpoints on healthcare priority setting. The results of this chapter indicated that participants' initial viewpoints partly remained stable over the course of the panel. In particular, participants' belief that access to healthcare is a basic human right and that priority setting should be based on patients' need for care and on the size and type of health gains remained important. Notable changes in participants' viewpoints concerned their increased support for healthcare priority setting, consideration of healthcare costs, and appreciation of cost-effectiveness of health technologies as a relevant decision criterion after participating in the panel.

This thesis contributes to the improvement of the decision-making framework by providing further insight into societal concerns for equity in healthcare priority setting, in particular for priority setting based on disease severity and the age of patients. The results of this thesis indicated that members of the public consider disease severity *and* the age of patients important in resource-allocation decisions in healthcare. This suggests that currently

applied decision-making frameworks (in the Netherlands as well as in other countries) may need to be adjusted in order to better reflect age-related preferences in society. Not only because these societal preferences are relatively strong but also because the age of patients is inextricably related to definitions of disease severity, like proportional shortfall. Moreover, the results of this thesis indicate that, in addition to adjusting the decision-making framework to (better) align the outcomes of decisions with societal preferences, decision makers could also consider to (better) align the process of resource-allocation decisions with societal preferences in order to increase the legitimacy of and public support for decisions. For example, by implementing citizens panels that inform or take on an even more prominent role in resource-allocation decisions.

Samenvatting

De vraag naar gezondheidszorg neemt snel toe, onder andere door de vergrijzing van de bevolking, de beschikbaarheid van steeds geavanceerdere en duurdere nieuwe behandelingen, en een hogere levensstandaard die de verwachtingen ten aanzien van gezondheid en gezondheidszorg doet toenemen. Omdat de middelen voor gezondheidszorg beperkt zijn, maakt de daaruit voortvloeiende druk op beschikbare budgetten prioritering in de gezondheidszorg onvermijdelijk.

Twee belangrijke doelstellingen van publiek gefinancierde zorgstelsels zijn het genereren van zoveel mogelijk (gezondheids) baten gegeven het zorgbudget en het nastreven van een eerlijke verdeling van gezondheid en gezondheidszorg. Om prioriteiten te stellen die bijdragen aan het bereiken van deze doelstellingen, integreert een toenemend aantal landen—waaronder Nederland—doelmatigheids- en rechtvaardigheidsoverwegingen in het besluitvormingskader. Bijvoorbeeld door rechtvaardigheidsoverwegingen mee te nemen in economische evaluaties van nieuwe behandelingen. Er zijn echter nog altijd belangrijke vragen over welke rechtvaardigheidsoverwegingen in aanmerking moeten komen, welk (relatief) gewicht deze overwegingen moeten krijgen, en hoe zij moeten worden meegenomen in vergoedingsbesluiten. In dit proefschrift zijn deze vragen aan de orde gesteld met als doel bij te dragen aan de verbetering van het besluitvormingskader door meer inzicht te verschaffen in maatschappelijke voorkeuren voor een rechtvaardige verdeling van de beschikbare middelen, in het bijzonder voor prioritering op basis van ziektelast en de leeftijd van de patiënten.

Hoofdstuk 2 onderzocht de normatieve rechtvaardiging en empirische ondersteuning voor de operationalisatie van rechtvaardigheidsoverwegingen in termen van proportional shortfall in Nederland. Proportional shortfall is afgeleid van de gerenommeerde severity en fair innings benaderingen en wordt gebruikt om ziektelast te definiëren en kwantificeren in vergoedingsbesluiten. De resultaten van dit hoofdstuk lieten zien dat het beslismodel, waarin op ziektelast gebaseerde rechtvaardigheidsoverwegingen worden meegenomen in economische evaluaties, breed wordt ondersteund en in toenemende mate wordt gebruikt voor het maken van vergoedingsbesluiten in Nederland. De resultaten lieten echter ook zien dat de theoretische en empirische ondersteuning voor het definiëren van ziektelast in termen van proportional shortfall beperkt is. Proportional shortfall houdt bijvoorbeeld onvoldoende rekening met leeftijdsgebonden voorkeuren in de samenleving, met name voor prioritering van jongere patiënten in vergoedingsbesluiten.

In de hoofdstukken 3 tot en met 5 is voortgebouwd op de resultaten van hoofdstuk 2 en zijn maatschappelijke voorkeuren voor prioritering in de gezondheidszorg op basis van ziektelast en de leeftijd van patiënten onderzocht. In hoofdstuk 3 worden de resultaten van een experiment gepresenteerd

waarin wij afzonderlijke en gecombineerde gewichten voor prioritering op basis van ziektelast en de leeftijd van patiënten hebben geschat. De resultaten lieten zien dat respondenten, gemiddeld genomen, een hoger gewicht gaven aan het vergoeden van behandelingen voor patiënten met een hogere ziektelast en voor jongere patiënten. De resultaten lieten verder zien dat de voorkeur voor prioritering van jongere patiënten relatief sterk was, omdat respondenten een hoger gewicht gaven aan jongere patiënten, zelfs wanneer zij een lagere ziektelast hadden dan oudere patiënten. In de hoofdstukken 4 en 5 zijn de resultaten gepresenteerd van een experiment waarin wij de betalingsbereidheid van respondenten voor gezondheidswinst in patiënten met verschillende leeftijden en niveaus van ziektelast hebben onderzocht, het laatste ook geoperationaliseerd in de context van terminale ziekte. De resultaten van deze hoofdstukken lieten zien dat de betalingsbereidheid, gemiddeld genomen, hoger was voor gezondheidswinst in jongere patiënten en in patiënten met een hogere ziektelast, zij het in een afnemende mate. De resultaten lieten verder zien dat de betalingsbereidheid voor gezondheidswinst in patiënten van verschillende leeftijden afhankelijk was van de ziektelast, de grootte van (en in mindere mate het type) gezondheidswinst, en de gevolgen van de ziekte. De betalingsbereidheid was, gemiddeld genomen, lager voor gezondheidswinst in terminaal zieke patiënten dan voor een vergelijkbare winst in patiënten die volledig van hun ziekte herstelden.

Hoofdstuk 6 onderzocht hoe verschillende visies op prioritering in de gezondheidszorg samenhingen met maatschappelijke doelmatigheids- en rechtvaardigheidsoverwegingen in vergoedingsbesluiten. De resultaten van dit hoofdstuk lieten zien dat de meerderheid van de respondenten tegen prioritering op basis van kenmerken van de patiënten, ziekte en behandeling was en dat deze groep van mening was dat prioritering uitsluitend gebaseerd zou moeten zijn op de zorgbehoefte van patiënten. Een minderheid van de respondenten zag in dat de beschikbare middelen schaars zijn en dat prioritering in de gezondheidszorg noodzakelijk is. Echter, slechts een klein aantal respondenten vond dat prioritering gebaseerd zou moeten zijn op kenmerken van de patiënten en de kosteneffectiviteit van behandelingen. De resultaten lieten zien dat respondenten die tegen prioritering waren minder geneigd waren om te prioriteren tussen concurrerende behandelingen en patiëntengroepen dan respondenten met een andere visie. Desalnietemin gaven de resultaten aan dat dit afhankelijk was van de specifieke besluitvormingscontext. De meeste respondenten waren inderdaad bereid om ten minste één keer de ene behandeling of patiëntengroep boven de andere te verkiezen, ongeacht hun visie op prioritering in de gezondheidszorg.

Hoofdstuk 7 onderzocht hoe deelname aan een deliberatief burgerforum van invloed was op de visie van de deelnemers op prioritering in de gezondheidszorg. De resultaten van dit hoofdstuk lieten zien dat de aanvankelijke visies van de deelnemers in de loop van het forum gedeeltelijk stabiel bleven. Met name de overtuiging van deelnemers dat toegang tot gezondheidszorg een

fundamenteel mensenrecht is en dat prioritering gebaseerd zou moeten zijn op de zorgbehoefte van patiënten en op de grootte en het type gezondheidswinst bleef belangrijk. Relevante veranderingen in de visies van de deelnemers hadden betrekking op de grotere steun voor prioritering in de gezondheidszorg, het in acht nemen van de kosten van zorg, en het besef van de relevantie van kosteneffectiviteit als beslis criterium na deelname aan het forum.

Dit proefschrift draagt bij aan de verbetering van het besluitvormingskader door meer inzicht te verschaffen in de rechtvaardigheidsoverwegingen die maatschappelijk relevant worden geacht in de prioritering van gezondheidszorg, in het bijzonder voor prioritering op basis van ziektelast en de de leeftijd van patiënten. De resultaten van dit proefschrift suggereren dat de bevolking bij prioritering zowel ziektelast als de leeftijd van de patiënten belangrijk vindt. Dit suggereert dat de momenteel gebruikte besluitvormingskaders (zowel in Nederland als in andere landen) mogelijk aangepast moeten worden om beter rekening te houden met leeftijdsgebonden voorkeuren in de samenleving. Niet alleen omdat deze maatschappelijke voorkeuren relatief sterk zijn, maar ook omdat de leeftijd van patiënten onlosmakelijk verbonden is met definities van ziektelast, zoals proportional shortfall. Bovendien wijzen de resultaten van dit proefschrift erop dat, naast het aanpassen van het besluitvormingskader om de uitkomsten van vergoedingsbesluiten (beter) af te stemmen op maatschappelijke voorkeuren, beleidsmakers ook kunnen overwegen om het besluitvormingsproces rond vergoedingsbesluiten (beter) af te stemmen op maatschappelijke voorkeuren om de legitimiteit van en publieke steun voor besluiten te vergroten. Bijvoorbeeld door het implementeren van burgerforums die besluiten informeren of zelfs een prominentere rol spelen in vergoedingsbesluiten.

PhD portfolio

Training

2019	Qualitative coding with Atlas.ti, Erasmus Graduate School of Social Sciences and Humanities, Rotterdam, the Netherlands
2018	Advanced choice modelling, Choice Modelling Centre, University of Leeds, London, United Kingdom
2018	Q methodology, Caledonian University, Glasgow, United Kingdom
2018	Self-presentation: Presenting yourself and your research, Erasmus Graduate School of Social Sciences and Humanities, Rotterdam, the Netherlands
2018	Great thinkers of the 20th century, Erasmus Graduate School of Social Sciences and Humanities, Rotterdam, the Netherlands
2017	Active learning, Risbo institute, Erasmus University Rotterdam, Rotterdam, the Netherlands
2017	Making an academic poster that stands out, Erasmus Graduate School of Social Sciences and Humanities, Rotterdam, the Netherlands
2017	Choice modelling and stated choice survey design, Choice Modelling Centre, University of Leeds, London, United Kingdom
2017	Patient Preferences in the delivery of healthcare, Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Rotterdam, the Netherlands
2017	Group dynamics, Risbo institute, Erasmus University Rotterdam, Rotterdam, the Netherlands
2017	Basic didactics and group dynamics for PhD students, Risbo institute, Erasmus University Rotterdam, Rotterdam, the Netherlands
2016	Foundations of Data Analysis Part I: Statistics Using R and Part II: Statistical Inference, University of Texas at Austin, Austin, United States of America
2016	Introduction to Bayesian Methods in Clinical Research, Erasmus University Medical Center, Netherlands Institute for Health Sciences, Rotterdam, the Netherlands
2016	Bayesian Statistics, Erasmus University Medical Center, Netherlands Institute for Health Sciences, Rotterdam, the Netherlands
2016	Quality of Life Measurement, Erasmus University Medical Centre, Netherlands Institute for Health Sciences, Rotterdam, the Netherlands
2016	Academic Writing for PhD Students, Erasmus Graduate School of Social Sciences and Humanities, Rotterdam, the Netherlands
2016	Scientific Writing in English for Publication, Erasmus University Medical Center, Netherlands Institute for Health Sciences, Rotterdam, the Netherlands

2016 Advanced Health Economic Modelling, Institute for Medical Technology Assessment, Erasmus University Rotterdam, the Netherlands

2016 Introduction to Discrete Choice Experiments, Institute for Medical Technology Assessment, Erasmus University Rotterdam, the Netherlands

Teaching

2019–2021 Rationing Health Care, bachelor programme in Health Policy & Management

2019–2021 Choices and dilemmas in healthcare, bachelor programme in Health Policy & Management

2018–2021 Pharmaceutical pricing and market access, master programmes in Health Economics, Policy and Law and Health Economics

2017–2021 Theses supervision, master programmes in Health Economics, Policy and Law and Health Economics

2017–2021 Healthcare Ethics, master programmes in Health Economics, Policy and Law and Health Care Management

2016–2021 Advanced Research Methods, master programmes in Health Economics, Policy and Law, Health Care Management, and European Master in Health Economics and Management

2016–2018 Statistics A and B, pre-master programme in Health Policy & Management

2016–2017 Theses supervision, bachelor programme in Health Policy & Management

2016–2017 Quantitative research in healthcare, pre-master programme in Health Policy & Management

2016–2017 Introduction to health sciences, bachelor programme in Health Policy & Management

Research grants and awards

2020 Research grant EuroQol Group, reference EQ Project 117-2020RA (co-investigator)

2019 Research grant EuroQol Group, reference EQ Project 20190920 (principal investigator)

2019 Research grant EuroQol Group, reference EQ Project 20190890 (co-investigator)

2019 Research grant EuroQol Group, reference EQ Project 20190860 (co-investigator)

2019 Dutch Society for Technology Assessment in Healthcare (NVTAG) award nomination for best paper published in 2018 by a young researcher in the NVTAG domain

2019 Fellowship of Postgraduate Medicine (FDM) award for best poster at European Cooperation on Healthcare conference, Rotterdam, the Netherlands

Invited presentations

2020 Equity weighting based on disease severity (oral presentation), seminar organised by Radboudumc Evidence Based Surgery Group, Nijmegen, the Netherlands

2019 Healthcare priority setting in the Netherlands (keynote speaker), symposium "Prioritisation according to disease severity – A barrier to treatment?" organised by Legemiddelindustrien (LMI), Oslo, Norway

Conference and symposium presentations

2021 Are Child QALYs Equivalent to Adult QALYs? (issue panel), Virtual International Society for Pharmacoeconomics and Outcomes Research (ISPOR) conference

2020 Willingness to pay for health-related quality of life gains in relation to disease severity and age (poster presentation), Virtual International Society for Pharmacoeconomics and Outcomes Research (ISPOR) conference

2020 Age-dependent willingness to pay for health-related quality of life gains in different equity contexts (oral presentation), Society for Medical Decision Making (SMDM) conference, Berlin, Germany (postponed to 2021)

2020 Severity and age dependent willingness to pay for quality of life gains (oral presentation), European Health Economics Association (EuHEA) conference, Oslo, Norway (postponed to 2022)

2019 Empirical evidence and future directions for equity weighting (oral presentation), symposium "Are all QALYs equal? Past, present and future of equity weighting", Rotterdam, the Netherlands

2019 Challenges in the design and implementation of coverage with evidence development schemes for medical devices (poster presentation), International Health Economics Association (iHEA) conference, Basel, Switzerland

2019 Challenges in the design and implementation of coverage with evidence development schemes for medical devices (poster presentation), European Cooperation on Healthcare conference, Rotterdam, the Netherlands

2018 How does participating in a citizens panel influence views on healthcare priority setting in the Netherlands (oral presentation), International Society on Priorities in Health (ISPH), Linköping, Sweden

2018 Looking back and moving forward: On the application of proportional shortfall in healthcare priority setting in the Netherlands (oral presentation), International Society on Priorities in Health (ISPH), Linköping, Sweden

2018 Does participating in a citizens panel influence views on healthcare priority setting in the Netherlands (poster presentation), European Health Economics Association (EuHEA) conference, Maastricht, the Netherlands

2018 Looking back and moving forward: On the application of proportional shortfall in healthcare priority setting in the Netherlands (oral presentation), European Health Economics Association (EuHEA) conference, Maastricht, the Netherlands

2017 Who should receive treatment? An empirical enquiry into the relationship between societal views and preferences concerning healthcare priority setting (oral presentation), International Health Economics Association (iHEA) conference, Boston, United Stated of America

2017 Who should receive treatment? An empirical enquiry into the relationship between societal views and preferences concerning healthcare priority setting (oral presentation with discussant), Lowlands Health Economics Study Group (LolaHESG) conference, Rotterdam, the Netherlands

Projects

2021-2022 Project on the social acceptance of cost-effectiveness analyses, funding from the National Health Care Institute (ZIN)

2019-2021 Projects focusing on adult preferences for health-related quality of life in children and adolescents, funding from the EuroQol Group

2018–2020 Pushing the Boundaries of Cost and Outcome Analysis of Medical Technologies (COMED), funding from the European Union's Horizon 2020 research and innovation programme

2018–2020 Broader Societal Benefits III (BSB-III) phase B, funding from a consortium of GlaxoSmithKline, AbbVie, and AstraZeneca

2017-2018 Burgerforum 'Keuzes in de Zorg', funding from the Netherlands Organisation for Scientific Research (NWO)

2016–2018 Broader Societal Benefits III (BSB-III) phase A, funding from a consortium of Pfizer, GlaxoSmithKline, AbbVie, Amgen, and AstraZeneca

Miscellaneous

2019–2021 Erasmus Centre for Health Economics Rotterdam (EsCHER): Coordinator

2019 Symposium "Are all QALYs equal? Past, present and future of equity weighting": Organiser

2018–2021 Academic Workplace Health Technology Assessment & Appraisal (Verzekerde Zorg), National Health Care Institute (ZIN): Member

2017–2021 Erasmus Choice Modelling Centre (ECMC): Member

2017–2021 Teaching activities of the department of Health Economics at Erasmus School of Health Policy & Management: Coordinator

2017 University Teaching Qualification, Risbo institute, Erasmus University Rotterdam, Rotterdam, the Netherlands

List of publications

Reckers-Droog VT, van Exel NJA, Brouwer WBF. Willingness to pay for health-related quality of life gains in relation to disease severity and the age of patients. Forthcoming in Value in Health. 2021.

Lipman SA, **Reckers-Droog VT**, Kreimeier S. Think of the children: A discussion of the rationale for and implications of the perspective used for EQ-5D-Y health state valuation. Forthcoming in Value in Health. 2021.

Rotteveel AH, **Reckers-Droog VT**, Lambooij MS, de Wit GA, van Exel NJA. Societal views in the Netherlands on active disinvestment of publicly funded healthcare interventions. Social Science & Medicine. 2021;272,113708.

Reckers-Droog VT, Jansen M, Bijlmakers L, Baltussen R, Brouwer WBF, van Exel NJA. How does participating in a deliberative citizens panel on healthcare priority setting influence the views of participants? Health Policy. 2020;124(2),143-151.

Reckers-Droog VT, Federici C, Brouwer WBF, Drummond MF. Challenges with coverage with evidence development schemes for medical devices: A systematic review. Health Policy and Technology. 2020;9(2),146-156.

Reckers-Droog VT, Goorden M, Kaminer Y, van Domburgh L, Brouwer WBF, Hakkaart-van Roijen L. Presentation and validation of the Abbreviated Self Completion Teen-Addiction Severity Index (ASC T-ASI): A preference-based measure for use in health-economic evaluations. PLoS One. 2020;15(9),e0238858.

Reckers-Droog VT, van Exel NJA, Brouwer WBF. Equity weights for priority setting in healthcare: Severity, age, or both? Value in Health. 2019;22(12),1441-1449.

Versteegh MM, Ramos IC, Buyukkaramikli NC, Ansaripour A, **Reckers-Droog VT**, Brouwer WBF. Severity-adjusted probability of being cost effective. PharmacoEconomics. 2019;37(9),1155-1163.

Reckers-Droog VT, Goorden M, Dijkgraaf M, van Eeren H, McCollister K, van Hakkaart- Roijen L. Unit costs of delinquent acts for use in economic evaluations. Journal of Mental Health Policy and Economics. 2019;22(2),71-79.

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.

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.

Baltussen R, Bijlmakers L, van der Burg S. van Dijk W, Groenewoud S, Jansen M, Tummers M, Helderman JK, Boer A, van Exel NJA, **Reckers-Droog VT**, Zwaap J. Draagvlak voor lastige keuzes: Eindrapport van het Burgerforum 'Keuzes in de zorg'. Nijmegen: Radboudumc. 2018.

About the author

Vivian Reckers-Droog is Assistant Professor at the department of Health Economics of Erasmus School of Health Policy & Management (ESHPM). Before starting her PhD trajectory in 2016, Vivian worked successively as a sociotherapist, team leader, and department head in an inpatient care facility for children and adolescents with complex mental health needs. Whilst being department head, she completed the pre-master programme in Health Policy & Management at ESHPM and the research-master programme in Health Sciences, specialisation Health Economics, at the National Institute for Health Sciences (NIHES). Vivian had previously completed bachelor programmes in Pedagogy and Fine Art.

Her research focuses primarily on societal concerns for equity in healthcare priority setting and the integration of such concerns with concerns for efficiency into the decision-making framework. Moreover, she conducts research on the valuation of child and adolescent health for use in economic evaluations of health technologies, on disinvestment decisions in healthcare, and on the challenges associated with coverage with evidence development schemes used to facilitate early access to new health technologies for patients.

Dankwoord

Toen ik een aantal jaren geleden startte met de pre-master Beleid & Management Gezondheidszorg om mij te professionaliseren in mijn toenmalige functie als leidinggevende in de jeugd-GGZ, had ik niet gedacht dat mijn loopbaan zo'n wending zou nemen. Ik ontdekte mijn passie voor het doen van onderzoek en besloot op basis daarvan om mij aan te melden voor de research master in Health Economics en niet zoals gepland voor de master Health Care Management. Ik heb mijn hart verder gevolgd en ben—nog voor het afronden van de research master—gestart met het promotietraject dat ik nu vol trots en dankbaarheid afrond.

Ik voel mij bevoorrecht met zoveel mensen om mij heen die mij de afgelopen jaren hebben geholpen en gesteund. Een aantal van deze mensen wil ik in het bijzonder bedanken.

Mariette, ik ben jou enorm dankbaar voor de ruimte die jij mij destijds hebt gegeven om het schijnbaar onmogelijke mogelijk te maken; het combineren van een intensieve baan met een intensieve opleiding. Die ruimte heeft mijn leven veranderd.

Werner en Job, ik wil jullie bedanken voor de fijne begeleiding tijdens mijn promotietraject. Ondanks (of misschien wel dankzij) jullie volle agenda's is het jullie steeds gelukt om mij op maat te begeleiden. Jullie gaven mij de ruimte om mij te ontwikkelen als zelfstandig onderzoeker, mijn interesses te verkennen en samenwerkingsrelaties met andere onderzoekers aan te gaan. Tegelijk waren jullie er wanneer ik jullie nodig had. Bedankt voor jullie altijd scherpe blik. Jullie zijn beiden een inspiratiebron voor mij en ik hoop dat ik jullie naam eer aandoe!

Daarnaast wil ik Rob Baltussen, Leon Bijlmakers en Maarten Jansen bedanken voor de uitnodiging om deel uit te maken van het Burgerforum project en voor de waardevolle bijdrage aan de publicatie waarop Hoofdstuk 7 van dit proefschrift is gebaseerd.

De leden van de promotiecommissie wil ik graag bedanken voor het beoordelen van mijn proefschrift en voor het opponeren tijdens mijn verdediging: Diana Delnoij, Johan Polder, Ardine de Wit, Jan van Busschbach, Maureen Rutten-van Mölken, Eddy Adang en Talitha Feenstra.

Mijn collega's bij Erasmus School of Health Policy & Management wil ik bedanken voor de fijne samenwerking de afgelopen jaren. Met name wil ik de PhDs bedanken met wie ik de afgelopen jaren het meest ben opgetrokken en heb samengewerkt: Judith, Klas, Meg, Stefan, Mariska, Linda, Sebastian H., Jannis, Marlies, Frédérique, Samare, Adriënne en tot slot Charlotte, inmiddels meer dan een collega. Ik kijk met veel plezier terug op de leuke tijd die wij samen hebben gehad! Lucas, ik wil jou ook met name bedanken. Ik geniet van de

hak-op-de-tak gesprekken waarin wij raadsels oplossen en ook steeds weer nieuwe raadsels bedenken. Ik leer veel van jou.

Tot slot wil ik mijn (schoon) familie en vrienden bedanken voor hun vertrouwen, steun en liefde. Martien en John, ik wil jullie enorm bedanken voor de mooie omslag en opmaak van mijn proefschrift. John, ik wil jou natuurlijk ook bedanken voor jouw begeleiding tijdens mijn verdediging. Ik kijk uit naar het moment dat ik mijn proefschrift mag verdedigen met mijn schoonvader en vader als paranimfen aan mijn zijde. Pap en mam, ik wil jullie in het bijzonder bedanken. Jullie hebben mij geleerd om hard te werken en altijd "meer te geven dan te nemen". Iedere dag zijn jullie daarin een voorbeeld voor mij. Sharon, jij hebt hetzelfde voorbeeld gehad en je bent misschien nog wel een hardere werker dan ik. Bedankt voor alle keren dat je tijd hebt vrijgemaakt om mee te lezen. Door jouw zorgvuldigheid wist je mijn Nederlandse teksten altijd te verbeteren. Tristan en Logan, het is leuk om je leven lang nieuwe dingen te leren. Ik hoop dat jullie nooit uitgeleerd raken en dat ik daarin een voorbeeld voor jullie ben.

Wouter, als laatste wil ik jou bedanken. Jij hebt mijn horizon verbreed en door jouw vertrouwen durf(de) ik mijn hart te volgen. Door wie jij bent, ben ik meer mijzelf.

ISBN 978-94-6416-455-8