Understanding severity:

Exploring public views on the meaning of severity of ill health

Mille Sofie Stenmarck, MD, BA (Honours)

Thesis submitted for the Degree of Philosophiae Doctor (PhD)

Faculty of Medicine
University of Oslo
Norway

October 2023

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo

ISBN 978-82-348-0418-2

Cover: UiO.
Print production: Graphic center, University of Oslo.

Table of contents

Acknowledgements

This thesis is part of the project Severity and priority setting in healthcare (SEVPRI) and is based on conversations with members of the Norwegian public. I would like to thank all participants for giving their time and for providing valuable and insightful reflections. This thesis would not have been possible without their contributions. I would also like to thank the Faculty of Medicine at the University of Oslo for giving me the opportunity to undertake this PhD project.

I am indebted to my supervisors for their continued support and guidance. To Mathias Barra, my main supervisor and Principal Investigator on SEVPRI, you have exceeded any expectations a PhD student could reasonably have of their supervisor. Thank you for sharing your experience and skills so generously, for always being available, and for making this PhD not only educative but also so enjoyable.

To Hilde Lurås, co-supervisor and department head at the Health Services Research Unit at Akershus University Hospital, where I have been stationed during this PhD. Thank you for your wise reflections and ongoing support, and for your words of encouragement throughout. To co-supervisors Rachel Baker and David GT Whitehurst, I am grateful for the time you have dedicated to providing valuable feedback on my work and for our workshops along the way. To Rachel, thank you for being my Q methodology tutor and for making every meeting an enjoyable one with your encouraging and positive attitude. To David, thank you for co-authoring all three studies in this thesis. The combination of your high-level insights and attention to detail made all the difference to our work.

I would like to thank my colleagues at the Health Services Research Unit. The level and breadth of research experience this community offers is surely unique, and has provided both a stimulating work environment and a supportive homebase. I would especially like to thank Jorun Rugkåsa, co-author and mentor in all things qualitative. To Tone Breines Simonsen, thank you for your valuable help in data collection. You and Nina Viksløkken Ødegaard have the answers to everything, really. I have been fortunate to share this experience with three other PhD students, Borgar Jølstad, Sindre Horn, and Marius Torjusen. Being your colleague and office mate has been a pleasure. I also want to thank colleagues from the clinic at Akershus University Hospital, and especially Grethe Kolloen, who made it possible for me to complete my junior doctor training while undertaking this PhD. Thank you for giving me the opportunity to be researcher and physician, and for believing that I could do both.

Throughout this PhD, I have been fortunate to be part of many inspiring seminars and research communities. I would like to thank colleagues and organisers of two separate seminars at the Brocher Foundation, which in different ways have contributed considerably to my work. To colleagues at SEVPRI's Swedish sibling, Just Severity, thank you for your thoughtful insights and enjoyable company throughout our annual seminars. To Daniel Wikler and Ole Frithjof Norheim, I am forever grateful for the opportunity to be part of the T.H. Chan School of Public Health at Harvard University. My time there was more inspiring and impactful than I can say. And to Roger Strand: with your perspective, honesty and sense of humour, you became my first teacher in the world of research, and continued as my teacher in life. You are a true mentor, and a true friend.

Last but not least, I would like to thank my family and friends. To my parents, thank you for your unwavering belief in me. To my fiancé, Herman, thank you for being my biggest champion and critic (in favourably unequal parts). You make me, quite simply, better and happier. And finally, to Bess. I am convinced you are the only grandmother to have exclaimed “oh no, what a waste!” to a grandchild's announcement of their medical school application. You dedicated your life's work to health policy and inspired me to follow suit, adamantly holding that efforts are best spent on systems-level change. I know you would have read this thesis with great pleasure and challenged me on every page of it. I hope it might have convinced you that medical school was, after all, not a complete waste of time.

Oslo, October 2023

Abbreviations

DSM	Diagnostic and Statistical Manual of Mental Disorders
HTA	Health technology assessment
ICD	International Statistical Classification of Diseases and Related Health Problems
NHS	National Health Service
NICE	The National Institute for Health and Care Excellence
NOU	Norwegian green papers (‘Norges offentlige utredninger’ in Norwegian)
QALY	Quality-adjusted life year
QM	Q methodology
Q2S	Q-to-survey
SEVPRI	Severity and priority setting in healthcare
UK	United Kingdom
USA	United States of America
VAS	Visual analogue scale
WHO	World Health Organization
WP1	Work Package 1
WP2	Work Package 2
WP3	Work Package 3

Summary

An increasing demand for healthcare resources makes resource scarcity a challenge for healthcare systems globally. As they struggle to meet the demands and expectations of the public, healthcare systems depend on priority-setting frameworks to guide resource allocation. The principles underlying these frameworks are the topic of ongoing debate in the academic literature, in policy, and in the public discourse. While many healthcare systems use principles of cost-effectiveness analysis to maximise healthcare delivery (however defined), these approaches are typically balanced against explicit ethical principles, guiding priority-setting decisions towards a fair and ethically defensible distribution (Daniels & Sabin, 2002; Ottersen et al., 2016). One of these ethical principles is *severity*. Applied as a priority-setting criterion in Norway, Sweden, the Netherlands, and the United Kingdom (Ministry of Health and Care Services, 2020; National Institute for Health and Care Excellence, 2022; Reckers-Droog et al., 2018; Riksdagsförvaltningen, 2018), severity is intended to modify cost-effectiveness analyses to ensure that treatment for illnesses that are considered particularly severe may be provided despite not necessarily meeting conventional cost-effectiveness thresholds. However, across these jurisdictions, severity is operationalised in different ways. The academic literature also points to ambiguity surrounding the meaning of severity, suggesting there is no consensus on the meaning of severity in a theoretical sense, either (Barra et al., 2020; Hausman, 2019a).

In recent years, there has been a participatory turn in health sciences and in public sector decision-making, with increasing emphasis on eliciting and including public views in policy-making (Baker et al., 2021). This is reflected in the many studies eliciting public views on severity in a healthcare context (Gu et al., 2015; Shah, 2009; Skedgel et al., 2022). While these studies demonstrate that there is widespread support among the public for severity as a relevant concern in priority setting, there is a paucity of research on what severity is taken to mean. Studies by Broqvist and colleagues and the Norwegian government-appointed Magnussen working group suggest that the citizenry associate various attributes with severity (Broqvist et al., 2018; Magnussen et al., 2015), but do not explore exactly what members of the public take severity to mean. The overall aim of this thesis is to address the knowledge gap surrounding public views on the meaning of severity. This aim is addressed via three objectives:

(i) to qualitatively explore public views on severity in a healthcare context via group interviews with members of the public;

(ii) to identify general viewpoints across different views on severity, by building on data from the group interviews and conducting a second round of data collection applying Q methodology; and

(iii) to examine the distribution of those viewpoints in a representative sample of the Norwegian population, also using Q methodological methods.

The abovementioned three objectives correspond to three separate studies. Study I offers a broad, explorative approach to identifying views on severity. Following group interviews with 59 members of the Norwegian general public, thematic analysis of interview transcripts resulted in the identification of a subjective, an objective, and a situational conceptualisation of severity. In Study II, statements pertaining to severity from the group interview transcripts were presented to a new set of participants. Study II is based on Q methodology, combining qualitative and quantitative techniques to provide an in-depth, exploratory approach to understanding complex concepts (Brown, 1993). A total of 34 participants individually rank ordered the selected statements, and these rank orderings were subsequently subjected to factor analysis to identify distinctive clusters of opinion. This resulted in four general viewpoints on severity: ‘natural lifespan’, ‘severity is subjective’, ‘objective measures and triage’, and ‘functioning and quality of life’. Building on findings from Study II, Study III provides an exploration of the prevalence and distribution of these viewpoints across a representative sample of the Norwegian population. The viewpoints were converted into vignettes and presented in a survey to 1174 Norwegians, recruited via quota sampling from an online panel delivered by a market research company. Findings from Study III suggest that no single vignette can be said to represent the citizenry’s views on the meaning of severity.

Overall, the findings presented in this thesis suggest there are multiple different views held by the public on the meaning of severity in a healthcare context. There seems to be a plurality of views, some of which are contrasting, and many that appear to represent additional concerns to those represented in policy operationalisations of severity. I discuss potential policy implications of this plurality of views, and consider the role of public views and how this plurality may be addressed in policy. I also consider the challenge facing policymakers considering the apparent misalignment between public views and policy.

Norsk sammendrag [Summary in Norwegian]

Det er økende etterspørsel etter helseressurser, som gjør ressursmangel til en utfordring for helsesystemer globalt. Mens helsesystemer strever for å møte etterspørselen og forventningene fra befolkning er de avhengige av prioriteringsretningslinjer for å veilede ressursfordeling. Prinsippene som ligger til grunn for disse retningslinjene er gjenstand for pågående debatt i både den akademiske litteraturen, i helsepolitisk sammenheng, og i den offentlige diskursen. Mange helsesystemer bruker prinsipper fra kost-effekt analyser for å maksimere helsetilbudet (definert på ulike måter), og disse er typisk veid opp mot eksplisitte etiske prinsipper som styrer prioriteringsavgjørelsene mot en rettferdig og etisk forsvarlig fordeling (Daniels & Sabin, 2002; Ottersen et al., 2016). Et av disse etiske prinsippene er *alvorlighet*. Det er anvendt som prioriteringskriterium i Norge, Sverige, Nederland, og Storbritannia (Ministry of Health and Care Services, 2020; National Institute for Health and Care Excellence, 2022; Reckers-Droog et al., 2018; Riksdagsförvaltningen, 2018) og er tiltenkt å modifisere kost-effekt analyser for å sikre at behandling for sykdommer som anses for å være spesielt alvorlige kan tilbys på tross av at de ikke nødvendigvis møter etablerte kost-effekt terskler. Men, det er mangel på konsensus rundt hva alvorlighet faktisk representerer på tvers av disse helsesystemene, og alvorlighet operasjonaliseres på ulike måter. Den akademiske litteraturen peker også på tvetydighet rundt hva alvorlighet betyr, som tyder på at det heller ikke er konsensus rundt hva alvorlighet betyr på et teoretisk plan (Barra et al., 2020; Hausman, 2019a).

De siste årene har det vært økt fokus på å inkludere befolkningssyn i politikkutforming i helsefeltet (Baker et al., 2021). Dette reflekteres i de mange studiene som utforsker befolkningssyn på alvorlighet i helsesammenheng (Gu et al., 2015; Shah, 2009; Skedgel et al., 2022). Selv om disse studiene demonstrerer at det er bred støtte blant befolkningen for alvorlighet som et viktig prinsipp i prioriteringsavgjørelser, mangler vi kunnskap om hva befolkningen mener at alvorlighet betyr. Studier av Broqvist og kollegaer og den regjerings-oppnevnte Magnussen gruppen viser at befolkingen assosierer alvorlighet med mange ulike attributter (Broqvist et al., 2018; Magnussen et al., 2015), men studiene utforsker ikke hva befolkningen mener at alvorlighet i seg selv betyr. Det overordnede målet med denne avhandlingen er å adressere dette kunnskapshullet ved å utforske befolkningens synspunkter på hva alvorlighet betyr. Det målet adresseres via tre delmål:

(i) ved å kvalitativt utforske offentlighetens syn på alvorlighet via gruppeintervjuer;

(ii) ved å identifisere generelle synspunkt på alvorlighet, basert på data fra gruppeintervjuene og ved å gjennomføre en ny runde med datainnsamling ved hjelp av Q metodikk; og

(iii) ved å studere hvordan de identifiserte synspunktene er fordelt blant et representativt utvalg av den norske befolkningen, også ved hjelp av Q metodologiske metoder.

De overnevnte delmålene korresponderer til tre ulike studier. Studie I har en bred, eksplorativ tilnærming for å identifisere ulike syn på alvorlighet. Gruppeintervjuer ble gjennomført med 59 deltakere fra den generelle befolkningen i Norge, og tematisk analyse av transkripsjonene fra gruppeintervjuene resulterte i tre konseptualiseringer av alvorlighet. Disse representerer en subjektiv, en objektiv, og en situasjonell konseptualisering. I Studie II ble sitater om alvorlighet hentet fra transkripsjonen av gruppeintervjuene og presentert til en ny gruppe deltakere. Studie II er basert på Q metodikk og kombinerer kvalitative og kvantitative metoder for å gi en detaljert, eksplorativ utforskning av komplekse konsepter (Brown, 1993). Totalt 34 deltakere ble bedt om å rangere sitatene ut ifra deres eget synspunkt på alvorlighet. Det ble deretter gjort faktoranalyse av disse rangeringene. Dette resulterte i fire generelle synspunkt på alvorlighet: 'naturlig livsløp', 'alvorlighet er subjektivt', 'objektive mål og triage', og 'funksjon og livskvalitet'. Studie III bygger på Studie II og utforsker prevalensen og fordelingen av disse synspunktene blant et representativt utvalg av den norske befolkningen. Synspunktene ble syntetisert til vignetter og presentert i en spørreundersøkelse til 1174 nordmenn, rekruttert via et kvoteutvalg fra panelet til en datainnsamler for markedsundersøkelser. Funnene i Studie III demonstrerer at ingen enkelt vignett kan sies å representere befolkningens syn på hva alvorlighet betyr.

Funnene som presenteres i denne avhandlingen demonstrerer at det, blant befolkningen, er mange ulike syn på hva alvorlighet betyr. Det ser ut til å være en pluralitet av ulike syn på alvorlighet, der noen syn står i kontrast til hverandre, og mange av disse synene representerer temaer som ikke er del av dagens operasjonaliseringer av alvorlighet. Jeg diskuterer potensielle implikasjoner av denne pluraliteten av syn, og vurderer rollen til befolkningssyn og hvordan denne pluraliteten kan adresseres. Jeg diskuterer også utfordringene som følger av et potensielt misforhold mellom befolkningssyn og helsepolitiske operasjonaliseringer.

List of studies

Study I

Stenmarck, M.S., Whitehurst, D.G.T., Lurås, H., Rugkåsa, J. *"It's hard to say anything definitive about what severity really is":* Lay conceptualisations of severity in a healthcare context.

[Under review at BMC Health Services Research, October 2023]

Study II

Stenmarck, M.S., Jølstad, B., Baker, R., Whitehurst, D.G.T., Barra, M. (2023). A severely fragmented concept: Uncovering citizens' subjective accounts of severity of illness. *Social Science and Medicine,* 330, 116046. https://doi.org/10.1016/j.socscimed.2023.116046

Study III

Stenmarck, M.S., Whitehurst, D.G.T., Baker, R., Barra, M. Charting public views on the meaning of illness severity.

[Submitted to Social Science and Medicine, October 2023]

1. Introduction

Healthcare resources are finite while the demand for healthcare is growing. This creates a dissonance between available healthcare resources and demand, and publicly-funded healthcare systems are forced to make difficult priority-setting decisions (Daniels & Sabin, 2002; Maynard & Bloor, 1998). Governments therefore rely on priority-setting principles by which resource distribution is determined, and on solid priority-setting frameworks to guide these decisions. Most publicly-funded healthcare systems seek to ensure both cost-effective and fair resource distribution (Hirose & Bognar, 2014), making priority setting both an economic matter and an ethical concern. Publicly-funded healthcare systems generally seek to balance the distribution of health resources in a way that safeguards two principles: (i) a cost-efficient allocation, where resources are directed where they are expected to have the largest effect, measured in some kind of health gain; and (ii) an ethically defensible distribution across the population, where resources are directed not only according to a logic of maximisation, but also according to the need for healthcare, measured according to various ethical quantifiers (Daniels & Sabin, 2002). One of these ethical quantifiers, across several jurisdictions, is *severity*.

In Norway, four separate government-appointed commissions have over as many decades worked to establish statutory principles for priority setting, presented in four separate green papers (‘Official Norwegian Reports’, in Norwegian abbreviated to ‘NOU’) (Ministry of Health and Care Services, 2020; NOU 1987: 23, 1987; NOU 1997: 18, 1997; NOU 2014: 12, 2014). Additional white papers produced by the Grund commission (NOU 1997: 7, 1997) and the Blankholm commission (NOU 2018: 16, 2018) also made important contributions to the establishment of the Norwegian priority-setting framework. *In toto*, these NOUs have resulted in three priority-setting criteria: health benefit, resources, and severity. The severity criterion is intended to modify cost-effectiveness analyses, ensuring that treatments for particularly severe illnesses may be afforded a higher priority than those for less severe illnesses. In recent years, severity has been implemented as a priority-setting principle in several other healthcare jurisdictions, including the United Kingdom (UK) (National Institute for Health and Care Excellence, 2022), the Netherlands (Reckers-Droog et al., 2018), and Sweden (Riksdagsförvaltningen, 2018). Despite the use of severity as a priority-setting principle across these healthcare systems, there is a lack of consensus on what severity actually means and these healthcare systems employ different operationalisations of severity. Severity also appears to be interpreted differently across different levels of the healthcare system, from the policy level to the clinic (Magnussen et al., 2015). Furthermore, in the academic literature, several different conceptualisations of severity are at play, and there is a lack of consensus on which interpretation of severity best represents its meaning in a healthcare setting (Barra et al., 2020; Olsen, 2013). Severity thus appears to be a complex and ambiguous concept.

There are several rationales for increasing public participation in policy development, including that soliciting the opinions of the population and implementing these in the policy-making process increases the democratic legitimacy of policies (Tenbensel, 2010). There is increasing emphasis on involving the public in priority setting in healthcare, which necessitates the elicitation of public views on priority-setting matters. Several studies exploring public views establish that members of the public consider severity to be a relevant concern in priority setting (Gu et al., 2015; Shah, 2009; Skedgel et al., 2022), but do not explore what members of the public take severity to mean. These studies employ different definitions of severity, but it is not evident or established that severity evokes the same intuitions across individuals. This leaves a paucity of research on what the public takes severity to mean.

The overall aim of this thesis is to identify and explore views on severity that exist among the Norwegian public, and to study the representation and distribution of those views across the population. The three studies included in this thesis address the knowledge gap on public views on severity in a healthcare context, by qualitatively examining data from group interviews across Norway (Study I); by exploring whether there are any discernible patterns across views expressed by members of the public that could represent general viewpoints on severity, using Q methodology and factor analysis with a new group of participants (Study II); and by analysing the distribution of each of those viewpoints across the population, using data from a cross-sectional survey distributed among a representative sample of the Norwegian population (Study III).

This thesis is structured as follows: in Chapter 2, I provide a background for priority setting in healthcare generally and to severity as a principle in priority setting specifically, and explore current knowledge of public views on severity. In Chapter 3, the aims and objectives of this thesis are outlined. In Chapter 4, the research methods of the three studies within this thesis are presented. The results of the three studies are presented in Chapter 5. In Chapter 6, I place the results of the three studies within the wider literature and present a broader discussion of the results within this context, considering also the wider policy implications of my findings. This is followed by a reflection on methodological considerations across the studies. Chapter 7 comprises a summary of the conclusions this thesis has led to.

2. Background

This chapter provides an outline of priority setting in healthcare generally and operationalisations of severity as a priority-setting principle specifically, both in Norway and in other publicly-funded healthcare systems applying a severity modifier. I then explore severity on a more conceptual level, considering how illness severity is understood in the literature as well as examining previous research on severity and public views. Finally, I outline the knowledge gap concerning public views on severity in a healthcare context.

2.1 Priority setting in healthcare

In recent decades there has been an increasing tension between medical possibilities and the need for economic restraint (Exter & Buijsen, 2012; Mechanic, 1997; Scheunemann & White, 2011; Williams et al., 2012a). In Norway, as in many other publicly-funded healthcare systems, there are greater challenges ahead, as the scarcity issues in healthcare become more pronounced (Hirose & Bognar, 2014; Magnussen et al., 2015; NOU 2023: 4, 2023). These increasing scarcity issues have several sources. Medical innovation and technical improvements in care have converted many previously deadly conditions into chronic ones. Rather than dying of such conditions, patients increasingly live with them—often for decades—and require continued medical support to do so. Furthermore, innovative research in areas such as genetics and oncology is leading to promising new therapies, but which are also costly and lead to rising consumer expectations of healthcare outcomes (Bjørnelv & Melberg, 2023). There is also a demographic explanation to increased scarcity in healthcare provision: the baby boomers of the 1940s to 1960s are transitioning into old age, placing greater demands on the healthcare sector—yet the continued population growth necessary to support the costs of the baby boomers' old age and declining health has not taken place (Hirose & Bognar, 2014; Mander, 2014). Finally, there is the inevitable dilemma of a publicly financed healthcare system: as the consumer is not paying for the goods they consume, they are likely to be less sensitive to costs, and inclined to demand more from the healthcare system (Hirose & Bognar, 2014). This requires that healthcare personnel balance providing the best possible care for their patients against their role as gatekeepers to the healthcare system. Finding an appropriate balance between these obligations is difficult. Altogether, these developments lead to a scenario where demand for healthcare is greater than supply, making healthcare resources limited (Daniels & Sabin, 2002; Scheunemann & White, 2011).

As the expectations of healthcare delivery increase (Ministry of Health and Care Services, 2019), national healthcare systems also struggle to provide the healthcare that the population expects. This is reflected in the increasing proportion of national budgets dedicated to healthcare, with healthcare spending set to outpace GDP growth in almost all OECD countries by 2030 (Centre for Medicare and Medicaid Services, 2021; Ministry of Health and Care Services, 2023a; OECD, 2019). However, although healthcare is an important and high-priority issue for the citizenry, there are other social goods that one can assume are also valuable, such as education, defence, infrastructure, and law and order (Scheunemann & White, 2011). Accordingly, there is a limit to how large a proportion of national resources can and should be directed towards healthcare. Healthcare systems must therefore set certain limits to care (Hirose & Bognar, 2014). This places heavy demands on healthcare systems to distribute resources effectively, and to direct resources where they are expected to have the greatest effect (Norheim et al., 2019). Priority setting becomes inevitable, from the policy level to the clinic.

Priority setting in healthcare is complex because the stakes are high: the consequences of priority-setting decisions have a direct effect on the healthcare goods available to the public. This is especially true of publicly-funded healthcare systems, where there are few or no alternatives to care available. Allocating resources to one group of patients inevitably incurs the rationing of resources from another (Butler, 1999). Priority setting involves trade-offs, creating dilemmas that set groups of patients up against each other, fighting for resources from the same, limited pool (Schoon & Chi, 2022). This scenario plays out almost daily on the front page of newspapers, as healthcare priority-setting decisions create ‘winners’ and ‘losers’ regarding access to care. This makes healthcare priority setting a contentious issue (Stenmarck et al., 2021) and places greater demands on the principles by which resource allocation is determined.

The terms ‘priority setting’ and ‘rationing’ tend to be used interchangeably in the literature (Williams et al., 2012a). Because ‘rationing’ specifically refers to the conscious limiting of medical care to restrain healthcare expenditure (Scheunemann & White, 2011), this thesis will henceforth apply the term ‘priority setting’, speaking more broadly to resource allocation situations in healthcare.

2.2 Ethical decision modifiers and severity

There is a broad field of research on the principles underpinning priority setting. Generally, health maximisation is a central principle in most healthcare systems, with the aim of maximising health benefits within set budgets. Health maximisation is commonly operationalised as cost-effectiveness: the costs and consequences of different therapies are examined and weighed against each other (Drummond et al., 2015), where the therapy that produces the greatest health gain for the lowest cost is preferable (Neumann et al., 2016).

However, distribution of healthcare resources is also an ethical concern. Health carries a great moral value because it is an important contributor to the range of opportunities open to the individual and their opportunity to live the life they wish to (Daniels, 2007). Priority-setting frameworks are generally intended to ensure not only maximisation of health gains across the population, but also fair distribution of those health gains (World Health Organization, 2010). In many publicly-funded healthcare systems, cost-effectiveness is therefore not the only consideration when allocating healthcare resources. While cost-effectiveness could be described as an implicit or explicit ethical principle, maximising health benefits for the population as a whole, there is also a concern for distributing resources fairly, beyond cost-effectiveness considerations. This makes acceptable allocation of goods an issue of balancing concerns for cost-effectiveness with explicit ethical principles (Pellegrino, 1986). For this reason, publicly-funded healthcare systems have, over decades, sought to establish processes of both cost-effective and fair resource allocation. As a result, many healthcare systems operate with ethical decision modifiers, intended to modify decisions guided by cost-effectiveness analysis. These ethical decision modifiers are intended to ensure that there is an ethical dimension to the distribution of available resources, and represent a willingness to sacrifice health gains to prioritise the worse off. This entails that the worse off (however defined) someone is, the stronger their claim (over other, less worse off individuals' claim) to a resource that benefits them.

There is no straightforward answer to who is worse off, and this is a heavily value-laden judgement. Ethical decision modifiers in healthcare priority setting come in many forms and are applied to varying degrees across different healthcare systems. Examples of such decision modifiers are special considerations for orphan drugs (Rosenberg-Yunger et al., 2011) and end-of-life (Cookson, 2013; Normand, 2012), the now-defunct Cancer Drugs Fund (NHS England Cancer Drugs Fund Team, 2016), and illness severity (Nord & Johansen, 2014).

Severity is implemented as a priority-setting principle in multiple healthcare systems, including in Norway (Magnussen et al., 2015), Sweden (Riksdagsförvaltningen, 2018), the Netherlands (Schurer et al., 2022), and, most recently, the UK (National Institute for Health and Care Excellence, 2022). Within these healthcare systems, severity is applied as a priority-setting principle or decision modifier (the terms are applied interchangeably) to ensure that resource allocation based on cost-effectiveness evaluations are modified according to the severity of the condition they are evaluating (Norheim, 2010). This allows for the approval of treatments that do not necessarily meet conventional cost-effectiveness thresholds, because the condition they are being considered as treatment for is considered particularly severe, and therefore deserving of prioritisation.

There is controversy surrounding all the above-mentioned decision modifiers, mainly because they allow treatments that exceed accepted cost-effectiveness thresholds to be approved to accommodate ethical concerns for distributing resources fairly (Clark & Weale, 2012). Severity is arguably particularly controversial, because there is no uniform agreement on exactly what ‘severe illness’ actually means (Barra et al., 2020). One of the consequences of discord on the meaning of severity is that its formulation into health policy varies between the systems that apply severity as a priority-setting principle. I will explore policy operationalisations of severity in Norway in Chapter 2.3.1 and in other jurisdictions in Chapter 2.3.2, but will first provide some context for the term ‘severity’. This thesis explores public views on severity in a Norwegian context, and severity in this thesis is translated from the Norwegian term ‘alvorlighet’. This is the most common translation between the terms in the literature. However, this is not necessarily a direct translation. The Norwegian term is used more commonly in everyday language than the English counterpart and is also more all-encompassing. ‘Serious’ is also a term used in the literature and may in some instances be a more apt translation of ‘alvorlighet’. In this thesis, I use the term ‘severity’ to describe different descriptions of ill health, including injury, illness, sickness, disease, and malady. I consider severity in terms of its meaning in a conceptual sense, rather than linguistically. I wish to explore how it is understood and conceptualised, and what it represents for members of the public in a healthcare context.

2.3 Severity as a priority-setting principle

As outlined above, severity is an important principle in priority setting, and severity is heavily entrenched in healthcare resource distribution in many priority-setting frameworks. However, severity is also applied broadly in the clinical setting and is an integral part of medical terminology—from conveying medical information to patients, to descriptions and definitions of diseases within professional medical guidelines. The World Health Organization's (WHO) International Statistical Classification of Diseases and Related Health Problems (ICD), arguably the most central tool for disease classification worldwide, consistently employs severity in descriptions of disease (WHO, 2022), as does the diagnostic manual used globally for mental disorder, the Diagnostic and Statistical Manual of Mental Disorders (DSM) (American Psychiatric Association, 2013). Finally, severity is a common, everyday term for members of the public and is a word most people have within their vocabulary. Severity is also used broadly in the public discourse on healthcare. Indeed, a search of the Norwegian newspaper database *Atekst* (Retriever Norway, 2023) using the keywords "illness severity" reveals 1,063 separate newspaper articles using the term in 2022 alone.

2.3.1 Severity in the Norwegian priority-setting framework

Severity is an important priority-setting principle; it is a central concept at the clinical level; and severity is a common, everyday term for the public and in the public discourse on healthcare. In the following section, I explore the meaning of severity in healthcare policy.

In Norway, as in other countries, priority setting in healthcare is based on two important principles: to maximise the overall health benefits to society (Wisløff, 2015), balanced against ensuring the right to equal access to healthcare (Husom, 2000). These principles are heavily entrenched in the healthcare system and demand the weighing of cost-effectiveness against a concern for fairness and equality. How best to operationalise these principles has been the topic of both public debate and rigorous theoretical and political work in recent decades, and Norway has a long tradition of systematically working with priority setting at the national level. This tradition originated in the 1980s with the first commission on priority setting in the healthcare system (Lønning et al., 1987). Since then, a total of five government-appointed commissions (Blankholm et al., 2018; Grund et al., 1997; Lønning et al., 1987, 1997; Norheim et al., 2014) and one government-appointed expert group (Magnussen et al., 2015) have assessed priority setting across the Norwegian healthcare sector, with emphasis on the context of secondary care, developing guidelines for priority setting from the policy-level to the clinic.

Severity has been central to the Norwegian priority-setting framework since its first formulation. The Lønning I commission proposed that priority setting should be based on the two criteria of severity of the disease and efficacy of the intervention (Lønning et al., 1987). A decade later, Lønning II proposed to add the criteria of cost-effectiveness, further recommended by the Grund commission (Grund et al., 1997). The three criteria of severity, efficacy, and cost-effectiveness were formally approved by the Norwegian parliament in 1999 (Ministry of Health and Care Services, 1999) and remain the foundation of the Norwegian priority-setting framework today. Subsequent commissions have discussed these criteria in-detail and developed new iterations and operationalisations building on the original formulations.

In 2013, the public discourse on healthcare priority setting became heated regarding the question of public reimbursement of the cancer drug *ipilimumab* (Yervoy) (Stenmarck et al., 2021). This culminated in the appointment of the Norheim commission, with a mandate to re-evaluate the current priority-setting criteria (Ottersen et al., 2016). The Norheim commission proposed three revised criteria: a health-benefit criterion, a resource criterion, and a health-loss criterion (Norheim et al., 2014). The latter was intended as a specification of the severity criterion from the Lønning II commission, but led to considerable public debate and discord when presented. It was controversial to many because it implied the weighting of past health loss, not only future health-loss as previous iterations were based on (Horn et al., 2021). The controversy surrounding this criterion led to the appointment of the Magnussen working group, who assessed the criterion and recommended an altered definition of severity (Magnussen et al., 2015). The version of severity proposed by the Magnussen working group, together with the health-benefit and the resource criteria, form the basis of the current priority-setting framework in Norway, presented in Table 1 (Norwegian Ministry & of Health and Care Services, 2017).

Table 1 The three criteria in the current Norwegian priority-setting framework.

Criterion	**Policy formulation**
Health benefit	The priority of an intervention increases with the expected health benefit.
Resources	The fewer resources an intervention requires, the greater the priority of that intervention.
Severity	An intervention's priority increases in keeping with the severity of the condition. The severity of a condition is determined by: - Risk of death or loss of function - The degree of physical or mental loss of function - Pain, physical or mental discomfort

In health economic terms, the two criteria of health benefit and resources translate into a cost-effectiveness calculation: the cost of an intervention (in the form of required resources) balanced against its expected effectiveness (the benefit of the intervention, operationalised as how effective the intervention is) (Ministry of Health and Care Services, 2020). The cost-effectiveness of a new healthcare intervention is summarised by the incremental cost-effectiveness ratio of that intervention, represented by the difference in cost between two interventions (e.g., the current intervention and a potential new intervention), divided by the difference in the effect of the two interventions (The Directorate of Health, 2012).

The health benefit criterion demands higher priority to interventions with a higher expected benefit, and is formulated in terms of ‘good life years’ (‘gode leveår’). This term is standardly operationalised via quality-adjusted life years (QALYs). The QALY is the most commonly used measure of health benefit in the health economic literature and is a representation of health state preferences (Salomon, 2017; Spencer et al., 2022; Weinstein et al., 2009). The QALY is intended to provide a quantitative measure of good life years, i.e., number of years lived in good health, and is based on the notion that health is a function of both length and quality of life (Prieto & Sacristán, 2003). The QALY-model stipulates that one year of life lived in ‘full health’ equals one QALY (Wichmann et al., 2020).

The resource criterion, centred on the costs of treatments, demands higher priority to interventions which require less resources. This is quantified by the amount of resources—such as costs of treatments, equipment, and health personnel hours—required by an intervention. While the health benefit and resource criteria thus balance the cost and effectiveness of a treatment, the severity criterion is intended to promote interventions which do not necessarily meet established cost-effectiveness thresholds, but which treat conditions considered to be particularly severe.

The three priority-setting criteria are intended to be applied throughout the healthcare sector. This means they are intended to apply both at the group level with health technology assessment (HTA) through to the clinical setting. The rationale for this is based on the view that these three concerns—severity, health benefit, and resources—are relevant both at the group level and at the clinical level (Magnussen et al., 2015). However, both the Norheim commission and the Magnussen working group acknowledged that decision-making at the group level and in a clinical setting require different approaches. At the group level, in HTA processes, decisions are made on behalf of (often large) patient groups, with consequential outcomes both economically and in terms of health. To ensure fair and consistent decision-making, this requires quantifiable methods. Likewise, it would be unreasonable to expect clinicians to quantify into HTA-comparable numbers the severity and the efficacy of every action they take regarding resource allocation in the clinic. To accommodate the needs of both HTA processes and healthcare personnel in the clinic, two versions of the health benefit criterion and the severity criterion were developed: a descriptive version for use at the clinical level and a quantitative version for use in HTA assessments at the group level (Ministry of Health and Care Services, 2020). The descriptive, clinical-level version contains the intuitions and aspects of the criterion which have a moral purpose in a priority-setting context (Magnussen et al., 2015).

The quantitative operationalisation of severity, intended for group-level decisions in HTA processes, is based on QALYs. In essence, the severity criterion modifies the health benefit and resource criteria, resulting in a priority-setting strategy based on severity-weighted cost-effectiveness. Severity is operationalised via weighted QALYs, meaning that QALYs gained for treatments of particularly severe illnesses are weighted heavier than QALYs for less severe illnesses. Greater value is placed on health improvements for Group A (with a more severe illness) than for Group B (with a less severe illness), despite the fact that the size of the health gain is equal. The weighting of QALYs according to the severity of a condition is typically done via choice-based population surveys, where respondents are presented with different trade-offs (Baker et al., 2010). This allows for exploration of which health states the citizenry consider to be more severe than others, and—via trade-offs—estimation of the relative willingness to trade longevity against improved health (Baker et al., 2010; Dolan et al., 2005).

There are multiple ways to account for severity within a QALY-based framework. In the current priority-setting framework, severity is operationalised (at the group level) as absolute QALY shortfall, referred to as absolute prognosis loss (in Norwegian: 'absolutt prognosetap'). This approach involves describing how many future QALYs one can expect to lose (compared to a standard reference life) as a result of a condition, at the time of diagnosis. The greater expected future QALY loss, the most severe the condition (Magnussen et al., 2015). This makes absolute QALY shortfall a measure of future loss of health, and is not backward-looking, as Norheim's health-loss criterion was (Ottersen et al., 2016). Higher priority is given to patient groups expected to lose more good life years due to illness over those expected to lose fewer good life years. Examples of conditions with high absolute QALY shortfall include childhood fatality and rheumatoid arthritis (Lindemark et al., 2014).

The descriptive operationalisation of severity is broad, and intentionally so. The rationale for this version is to enable decision-making at lower levels of the health sector, making room for judgement on the part of the clinician who knows the patient and their situation. The Magnussen working group expressed that there should be room for judgement in special circumstances, with a definition of severity that encompasses aspects of severity that are important at the clinical level, and which allows the clinician to take the entire situation of the patient into consideration. The Magnussen working group's report provides examples of such circumstances, including the role of dignity, whether the condition might lead to decreased functioning, whether the patient has demanding parental responsibilities, and how a condition might affect the ability to work or engage with the community (Magnussen et al., 2015).

2.3.2 Operationalisations of severity in other jurisdictions

Other countries have engaged in similar processes to Norway in developing healthcare priority-setting principles. Healthcare systems in the Netherlands, Sweden, and the UK also employ severity as a decision modifier, and have to varying degrees implemented severity in priority-setting frameworks. However, these healthcare systems interpret severity in different ways than in the Norwegian model and employ different operationalisations of it. In the Netherlands, proportional QALY shortfall is employed (Reckers-Droog et al., 2018). This entails considering how many QALYs one can expect to lose due to a condition, relative to how many remaining QALYs one could expect to have in one's lifetime if not afflicted with that condition. The National Healthcare Service (NHS) in the UK, which implemented severity as a decision modifier in 2022, applies both absolute and proportional QALY shortfall in its approach to severity (National Institute for Health and Care Excellence, 2022). In the Swedish priority-setting framework, severity is operationalised according to severity levels—'low', 'moderate', or 'high' severity—along the two main axes of current and future ill health (Prioriteringscentrum, 2017; Riksdagsförvaltningen, 2018). Within these axes, severity is measured according to impairment of bodily functions, activity limitations, participation restrictions, occurrence, duration, risk of future ill health, risk of premature death, and time to death (Broqvist, 2018; Calltorp, 1999). Furthermore, both France and Belgium employ severity to determine reimbursement evaluations and price determinations for pharmaceutical interventions, via evaluating levels of severity. However, neither of these jurisdictions have any specific definition of severity (Franken et al., 2015; Magnussen et al., 2015).

2.4 The meaning of severity

This section provides an overview of previous research on severity in two different spheres: within the academic literature on illness severity itself, and within empirical work on public views on severity. The section on public views contains a consideration of the role of public views in priority setting and decision-making, as well as an overview of previous research on public views and severity.

2.4.1 The literature on severity

In the setting of healthcare distribution, the notion of severity is drawn from different theories of distributive justice. These theories emphasise the role of the worse off and the legitimate claims to healthcare resources the worse off have (Daniels, 2007; Rawls, 1999) and imply that prioritising those considered to be worse off, interpreted in many jurisdictions to be represented by the most severely ill, has high social value (Brock, 2002). However, across the various healthcare literatures invoking the concept of severity, there is no universal agreement on exactly what is meant by it, and severity is considered an ambiguous concept (Barra et al., 2020; Hausman, 2019b; Shah, 2009).

Olsen argues that there are various understandings of severity at play, and that (at least) four different categories of severity can be drawn from the literature: severity determined by (i) the degree of poor health, (ii) short remaining lifetime, (iii) poor prognosis, or (iv) the size of the health loss (Olsen, 2013). In the first approach, which Olsen argues is most commonly applied in the literature, an improvement of health state from a relatively low (severe) level has greater social value than an improvement of health state in a less severe illness. Nord's approach to severity is reminiscent of this, and he has since developed more work determining illness severity based on a severity scale, measuring severity in terms of health-related quality of life (Nord, 1993b). The second approach outlined by Olsen is based on remaining years of life without treatment, and is reminiscent of end-of-life arguments that have been prominent in the work of UK health economists (Kinghorn & Coast, 2019; Pinto-Prades et al., 2014; Shah et al., 2018) and, until recently, in NICE guidelines (NICE, 2009). Studies such as Hansson and colleagues' also appear to subscribe to this approach, using life expectancy as an indicator of illness severity (Hansson et al., 1994).

The third approach is centred on prognosis, suggesting that the poorer the prognosis, in the form of relatively fewer remaining life years due to illness, the more severe the illness, and the higher priority that patient group is given. While the third approach also compares to end-of-life arguments, this approach takes both remaining years of life and health state into consideration. As such, the third approach combines the first and second approaches. Prognosis, in an isolated sense, is reminiscent of the principle of rule of rescue (Cookson et al., 2008, 2008), though the latter typically involves identifiable individuals rather than hypothetical future individuals or statistical lives. The fourth approach represents a departure from other operationalisations of severity, considering the net expected health loss, conferring priority according to the size of the estimated health loss compared to a reference life. As Olsen points out, this approach bears comparison to fair-innings arguments, centred on the aim of reducing overall inequalities in health outcomes and based on the argument that individuals have the right to a certain quality-adjusted life expectancy (Williams, 1997). Severity understood as net expected health loss also compares to the notion of prioritarianism (Daniels, 2007; Parfit, 1991) and takes past, present, and future health loss into consideration (Ottersen et al., 2014). The emphasis on net health loss involves an age-centred view of severity, with conditions affecting younger individuals necessarily considered more severe than conditions affecting the old (other things being equal), as the net health loss of those who have already reached a high age will always be smaller.

Evidently, there is ambiguity on the different meanings and interpretations of severity in the health economics literature, and the different approaches to severity lead to diverging HTA outcomes and policy operationalisations of severity, as illustrated above. The ambiguity surrounding severity is also evident in the literature relating to clinical use of the term. I highlighted above the use of severity in clinical guidelines, but severity is also applied in the wider literature surrounding different clinical specialties. For example, Zimmerman and colleagues point to the role of severity in the field of mental health, emphasising how severity influences decisions about level and scope of care (Zimmerman et al., 2018). They highlight the uncertainty surrounding the term, and the lack of consensus on what constitutes severe illness in a mental health setting. Severity is also extensively applied in the field of genomics, invoked as a form of filtering system for access to therapies and services, but without consensus on what defines a severe genetic disorder (Boardman & Clark, 2022). The ambiguity surrounding severity also became evident during the SARS-CoV-2 pandemic, when applying the standard principles for priority setting in a public health context proved difficult, especially in terms of priority setting and vaccine distribution as a preventative healthcare measure (Horn et al., 2021). The literature also demonstrates ambiguity on how healthcare professionals interpret severity. The Magnussen working group distributed a survey among Norwegian policymakers, healthcare workers, and patient organisations, and while they do not provide details of their findings, they write that widely disparate definitions of severity emerged, with no consensus on the meaning of severity or on how it ought to be applied in a priority-setting context (Magnussen et al., 2015).

2.4.2 Public views and severity

The role of public views

The principles determining how goods are allocated are telling of the values that underpin a society. This requires that policymakers have an awareness of the underlying views and values on which the choices of distribution of public goods are based (Hirose & Bognar, 2014), and that decisions on resource distribution are aligned with the social values of the citizenry they represent (Norheim et al., 2014). Eliciting these views requires active engagement with the public. Across the literature on both healthcare priority setting and on the role of public participation in healthcare, there exists a wide range of conceptualisations of who constitutes ‘the public’, and who can reasonably be considered to represent the public in policy processes.

‘The public’ is commonly evoked in reference to a given population, but can represent a number of different groups of people and different sub-communities within a larger setting (Stewart, 2016). As such, there is no one, single public (McHugh, 2022). In terms of public views on healthcare provision, ‘the public’ can refer to a number of different subsets of the public, including groups within the health services, patient advocacy groups, and elected representatives (McHugh, 2022).

In this thesis and in the research reviewed below, ‘the public’ refers not to subsets of the public (either within or beyond the healthcare system), but to the wider citizenry. This is the most common construction of the public in the literature on healthcare policy (Tenbensel, 2010). I will also use the terms ‘the public’ and ‘the citizenry’ interchangeably. While these terms carry different meanings in a judicial setting, such distinctions are not relevant to the field of healthcare and the terms can therefore be applied to represent the same thing: all members of the public. This may include patients and healthcare professionals, who—independently of their health state or profession—are also members of the public.

There are several rationales for greater public participation in health policy. One such rationale builds on the notion of democratic legitimacy (Tenbensel, 2010) and that, when conducted in a democratic context, health policy processes depend on public participation to ensure accountability and transparency. In democracies, it is the right of citizens to expect and demand that choices that bear a great consequence on their lives, such as healthcare priority-setting decisions, are made in a transparent and accountable way (Hirose & Bognar, 2014). To ensure the democratic legitimacy of policies, the public should either be able to recognise policies as representative of their own views, or—by compromise—that public views are taken into account as part of the policy-making process (Barra et al., 2020).

This sentiment is reflected in the accountability for reasonableness framework, describing the conditions of a fair decision-making process (Daniels, 2000). One of the principles behind this framework is that, because societies struggle to agree on the principles underlying priority setting, it is important that the processes by which priority-setting decisions are made are fair.

Ensuring fairness involves transparency in the decision-making process (Daniels & Sabin, 1998), making transparency the hallmark of a fair process (Daniels, 2000; Martin et al., 2002). The public should have knowledge of the rationales behind priority-setting decisions and the opportunity to gain an understanding of the grounds of those decisions. Across the political spectrum, there is also broad consensus that policy should not only be transparent, but—to protect democratic ideals—citizen participation ought to be encouraged as an 'intrinsic democratic good' (Stewart, 2016).

As such, public participation via eliciting the views of citizens arguably represents a democratic good in and of itself. This can be seen as one rationale for public participation. Well-functioning democracies therefore ought to ensure that the public has a voice by fostering active citizenship and ensuring that the views of the public are actively sought out and listened to (Broqvist et al., 2018; Frankish et al., 2002; Litva et al., 2002). Priority-setting decisions should reflect the views and norms of society, and it is therefore critical that these are explored and known to policymakers (Williams et al., 2012a). To retain their legitimacy, organisations within the public sector are dependent on public participation. When policies involve the use of contested, complex terms, such as severity, eliciting the views of the public is arguably even more important.

A second rationale for public participation in policy development and priority setting is that public participation can produce better policy decisions (Stevenson, 2016). Participation has instrumental value and makes an important 'contribution to policy-relevant knowledge' (Tenbensel, 2010). This second rationale rests on the assumption that exploring and implementing public views improves the quality of priority-setting decisions by providing relevant and valuable knowledge (Broqvist, 2018; Frankish et al., 2002; Mitton et al., 2009). Because healthcare is an important good, there is an increasing involvement of public views in healthcare policy-making (Mullen, 1999) and it is no longer seen as adequate for decisionmakers to 'act as the sole conduit for public knowledge' (Stewart, 2016). To develop solid, legitimate policies, policymakers therefore depend on knowledge of key views and values within the society they serve. To elicit these views, they rely on the public's expertise by exploring the normative assumptions surrounding a given topic (Lehoux et al., 2009).

A third rationale for the elicitation and involvement of public views in policy is associated with the consequence of healthcare priority-setting outcomes, and that citizens have a vested interest in how healthcare resources are rationed and allocated. The distribution of healthcare resources within a public healthcare system has considerable bearings on the lives of the citizens within that public. It seems reasonable that the views of the citizenry ought to play some role in the formulation of policies which have such consequential impacts on it (Gustavsson & Lindblom, 2023). That the consequences of priority setting are important to the public is a rationale in itself for eliciting and incorporating public views in policy-making (Reich, 1988). Public preferences should therefore play a role in decision-making surrounding which principles underlie the distribution of scarce public resources (Schwappach, 2002), and citizens ought to have the opportunity to contribute to discussions on these matters. The inclusion of public preferences is particularly important in publicly-funded healthcare systems, where healthcare provision is financed by taxes. In a system where the citizenry are both the payers of healthcare provision and the beneficiaries of it, it is important that their views play a part in informing policies (Culyer, 2017; Morrell et al., 2018).

Evidently, the voice and views of the public have value in the context of priority setting in healthcare, and there is a rich literature to support this (Brock, 2002b; Gustavsson & Lindblom, 2023; Powers, 2005). This is also reflected in the participatory turn witnessed in recent years in health sciences, including in the field of priority setting, towards increased emphasis on citizen participation in policy development (Abelson et al., 2013; Baker et al., 2021; Baltussen et al., 2017). How public participation can be ensured is a complex matter in itself, and the degree to which public views should be implemented in decision-making is undefined (Litva et al., 2002). This also varies between countries and different healthcare systems, and between the different levels of healthcare provision within healthcare systems. This matter is beyond the scope of the thesis, and I will therefore conclude this section by stating that public views are central to ensuring transparent and legitimate priority-setting processes, making knowledge of public views on priority-setting principles of value.

Public views on severity

There are multiple studies exploring public views on the relevance of severity as a decision modifier (Gu et al., 2015; Jacobsson et al., 2005; Nord, 1993; Nord & Johansen, 2014; Schwappach, 2002; Shah, 2009; Skedgel et al., 2022). Shah conducted a systematic review on the published literature concerning public views on severity in the context of economic evaluation and priority setting (Shah, 2009). This included 21 empirical studies on severity-related preferences, most of which involved interviews and surveys with members of the general public. Gu and colleagues also explored the public's views on how priority should be given in healthcare, conducting a literature review on factors which members of the public believe should confer priority (Gu et al., 2015). In their review of 64 studies, Gu and colleagues identified nineteen which elicited preferences on severity, and remarked on the heterogeneity in the provided definitions and that some do not define severity at all. The studies included in the literature reviews by Shah and Gu and colleagues varied from small-sample group discussions to large population surveys. They were conducted across various geographical locations, including Denmark, Norway, Iceland, Sweden, Finland, Australia, South Korea, the UK, and the USA.

Both review papers concluded that there is widespread support for severity as a relevant concern in priority setting. Multiple studies within the review papers demonstrate that illness severity is considered relevant for priority setting, with members of the public in different ways expressing a preference for prioritising the severely ill. Gyrd-Hansen demonstrates a preference for prioritising treatment to the more severely ill over those with less severe illness, defining severity as health benefit (Gyrd-Hansen, 2004). Whitty and colleagues explore multiple public elicitation studies on healthcare priority setting, some of which concern severity; they demonstrate that, overall, there is support for prioritising the severely ill, but across these studies severity is defined both according to health state, health gain, quality of life, age at onset, etc. (Whitty et al., 2014). Shiroiwa and colleagues demonstrate a higher willingness to pay for severe illnesses (understood by health state) in a survey study (Shiroiwa et al., 2013), and Oddsson demonstrates a reluctancy among the citizenry to allocate only limited resources to the severely ill (measured via severity levels, but without providing a definition of severity) (Oddsson, 2003). Data from two different surveys, one with economics students and one with prospective jurors, demonstrate a preference for treating the severely ill, though neither define severity (Ubel et al., 1996, 1998). An interview study with employees at the Norwegian National Institute of Public Health makes similar findings, determining severity according to seven levels of physical disability (Nord, 1993b). In a survey study with health politicians, Nord demonstrates a preference for at least equal priority to the severely ill (Nord, 1993a). Considering public views on the role of illness severity as a specific criterion in priority-setting frameworks, a citizen council in The National Institute for Health and Care Excellence (NICE) tasked with deliberating severity judged that severity ought to be a priority-setting criterion, but do not describe what is meant by severity (NICE Citizens Council, 2008). Two general public survey studies suggest severity is considered one of the most relevant concerns in resource allocation, with Winkelhage and Diederich defining severity according to need and rule of rescue (Winkelhage & Diederich, 2012), and Ryynänen and colleagues leaving severity undefined (Ryynänen et al., 1999). Other general public survey studies demonstrate that members of the public may even consider severity to be the most important concern in resource allocation, but none of these studies describe what severity is interpreted to mean (Diederich et al., 2012; Dolan & Shaw, 2003; Green, 2009; Lim et al., 2012; Linley & Hughes, 2013).

Since the publication of the reviews by Shah and Gu and colleagues, several studies have continued to explore public views on severity. Richardson and colleagues test the ‘severity hypothesis’ (measuring severity by increase in a patient’s utility), establishing that health gains to those with relatively more severe illness are considered relatively more valuable (Richardson et al., 2017). Nord and Johansen build on Shah’s review, selecting 17 studies to explore the notion of a quantifiable ‘severity gradient’, representing a trade-off between prioritising treatment for the severely ill (understood in QALY terms) and maximising health gains (Nord & Johansen, 2014). They conclude that there is good evidence that severity matters, but that the degree to which it matters is widely varying.

Overall, these studies demonstrate that severity is considered a central issue in healthcare priority setting for members of the public across many geographical locations, and in various ways explore how members of the public think resource allocation decisions ought to be made in light of a concern for the more severely ill. However, none of these studies appear to explore the meaning of severity itself. They provide varying conceptions or definitions of severity, sometimes provide them for the reader but not participants, and sometimes provide no definition at all. It is also not clear whether participants across the studies were aware that they were answering questions intended to assess the relevance of severity, as some of these studies do not appear to have used the term in data collection. When definitions are provided, and the authors describe what they mean by severity when asking members of the public to consider its relevance to priority setting, a wide array of definitions are at play, including capacity to benefit from treatment, size of health loss, quality of life if untreated, and prognosis if untreated (Gu et al., 2015; Shah, 2009; Skedgel et al., 2022).

These definitions mostly involve some form of QALY shortfall, representing different measurements of loss of quality and quantity of life, but without exploring if there are other potential attributes of severity that might be considered relevant to the public. While an exploration of the meaning of severity was not the aim of the studies reviewed by Shah and Gu and colleagues, neither the reviews nor the studies included in the reviews problematise that varying definitions are applied across the literature. Overall, these studies and reviews appear to assume that members of the public have similar intuitions about what severity means, and that when they ask the public about severity via the various definitions they use, they assume respondents are indeed providing viewpoints about severity. However, without exploring what severity is interpreted to mean, and being consistent in which definition of severity participants are asked to consider, one cannot know if severity means the same thing to different people, or if these public elicitation studies are measuring the same thing.

Current parliamentary guidelines on priority setting in Norway state that “today's [absolute QALY shortfall operationalisation of severity] provides a reasonable expression of society's emphasis on severity” (Ministry of Health and Care Services, 2016). To my knowledge, there are no studies, either in a Norwegian or an international context, that present findings to support this claim. While multiple studies demonstrate that the public consider severity to be relevant to healthcare priority setting, little has been done to establish what members of the public associate with the term, or to explore what severity is taken to mean. The only study to have explored public views on the meaning of severity is by Broqvist and colleagues (Broqvist et al., 2018). In a qualitative interview study with members of the Swedish public, they explore what Swedish citizens consider relevant in evaluating the different levels of severity in the Swedish priority-setting system. The authors demonstrate that severity, according to members of the Swedish public, is a multifactorial concept. If the views of the Swedish citizens within their study are representative of those held by other citizenries, and the findings are transferable to other contexts, public views of severity may not align with operationalisations of severity incorporated in other healthcare jurisdictions.

Although Broqvist and colleagues make a valuable contribution to understandings of public views on severity, the study is centred on public views on severity levels, asking about their views on the different aspects of the Swedish Severity Framework and to what degree they consider the framework relevant to evaluating severity levels (Broqvist et al., 2018). While their study provides insight into views on severity as they pertain to the Swedish framework, important questions remain about what severity, in a broader sense and preceding policy formulations of the concept, is taken to mean. As such, there is a paucity of research on public views on the meaning of severity.

2.5 The knowledge gap

Severity is a central principle in multiple healthcare jurisdictions and is commonly applied in decision-making at the clinical level. Yet, despite the substantial health-related literature invoking the concept of severity, the meaning of severity itself has not been extensively explored. There is a lack of consensus on any theory of what is meant by ‘severity’ in the healthcare, and there is no agreement on which operationalisation of severity best facilitates a concern for the most severely ill. Furthermore, there is a lack of knowledge on public views regarding the meaning of severity (Barra et al., 2020). While there appears to be widespread public support for severity as a relevant concern in healthcare priority setting, there is a paucity of research both on what severity is taken to mean by members of the public, and on how views on severity are distributed among the population. It appears that severity as a priority-setting principle is applied widely, yet explored narrowly. The ambiguity surrounding the meaning of severity suggests that severity remains a controversial concept (Hausman, 2019b).

This thesis addresses the paucity of research on public views on *the meaning of severity*, seeking to explore, summarise, and quantify the views of the public on this issue. Because severity is a central principle in healthcare priority setting, the way severity is interpreted has a direct consequence on how resources are prioritised (Olsen, 2013). In a public healthcare system, determining severity is part of the equation on *who* gets *what*. However, the transparency of priority-setting frameworks remains clouded when underlying principles of that framework are not fully accounted for. Lack of knowledge on public views surrounding the meaning of severity in a priority-setting context arguably undermines the democratic legitimacy of priority-setting decisions that invoke the term. Increased knowledge on public views surrounding severity would enhance policymakers’ ability to be responsive to public views and preferences, and better enable them to make decisions in line with public views.

The findings presented in this thesis can provide policymakers—both in Norway and in other countries applying illness severity as a principle in healthcare priority-setting frameworks—with increased understanding of public views on severity in a healthcare context, as well as knowledge about the extent to which different views are supported in the population.

3. Aims

3.1 Severity in priority setting in healthcare (SEVPRI)

This thesis is part of the Severity in priority setting in healthcare (SEVPRI) research project. SEVPRI was designed to address the knowledge gap on the meaning of severity in the context of health and illness. The overall aim of SEVPRI is to improve on the way severity is discussed in the public discourse, by better understanding views on severity held by members of the public and exploring how people's values impact their views on severity. By seeking to identify and understand members of the public's views on severity, and by quantifying how those views are distributed across the Norwegian population, the project aims to help in achieving a less contentious public discourse on priority setting.

SEVPRI has three main aims, which are explored via three separate work packages. This thesis, comprising Work Package 1 (WP1), is named 'Subjective accounts of severity'. The idea behind WP1 is to ascertain and characterise the public's views on severity in a healthcare context, with both qualitative and quantitative components. Work Package 2 (WP2) is concerned with severity in relation to fair distribution of health outcomes. In Work Package 3 (WP3), the subjective accounts of severity identified in WP1 are mapped onto theories.

3.2 Aims of the thesis

The overall aim of this thesis is to investigate members of the Norwegian public's views about severity in a healthcare context. The specific objectives of the different studies within this thesis, directed towards addressing the overall aim of ascertaining and characterising views on the meaning of severity, are listed below. Each objective corresponds to one of the three studies presented in this thesis:

1. To qualitatively explore public views on the meaning of severity in a healthcare context (Study I). This objective is addressed in a qualitative study, based on data collected from group interviews with members of the public.

2. To identify patterns across different viewpoints on the meaning of severity held by the public and identify potential areas of consensus and contention (Study II). This

objective is addressed via a mixed-methods study using data collected from rank ordering exercises completed by a different sample of general public participants.

3. To explore the distribution of the viewpoints identified in Study II across the wider population, describing the representation of each of the viewpoints and how they are distributed across the Norwegian population (Study III). This objective is addressed in a cross-sectional survey with a representative sample of the Norwegian population.

The results of Study II also form the basis of a separate study, which is part of WP3, and where I am a co-author (Jølstad et al., *forthcoming*). That paper does not form part of this thesis.

3.3 Ethical approvals and funding

Both SEVPRI and this PhD project are funded by the Norwegian Research Council (project number 303724), at Akershus University Hospital. As a PhD student, I have been enrolled at the Faculty of Medicine at the University of Oslo. The Regional Committee for Medical and Health Research Ethics deemed the SEVPRI study outside their mandate and the remit of the Norwegian Health Research Act (ref. no. 186284). Accordingly, the Data Privacy Officer at Akershus University Hospital advised that data collection for SEVPRI could be conducted, following a detailed Data Protection Impact Assessment (PVO. Nos 20_200 and 21_200). Akershus University Hospital and the Principal Investigator, Mathias Barra, are responsible for project oversight, including all aspects of ethical research conduct and data privacy. Within the studies presented and this thesis overall, the names of people and organisations have been either deleted or anonymised. All participants in the project gave their written and/or oral consent.

Throughout the research process, I have followed several ethical guidelines to ensure sound ethical conduct. These include the Guidelines for Research Ethics in the Social Sciences and the Humanities (NESH, 2021) and the Guidelines for Research Ethics and Scientific Assessment of Qualitative Research Projects in Medicine and Healthcare (NEM, 2009). Both I, co-authors, and the SEVPRI research team considered and collaboratively discussed ethical concerns during the design, data collection, and analysis of the various studies presented in this thesis, as described in-detail in Chapter 6.7. Storage and handling of data was done according to guidelines of the Data Privacy Officer at Akershus University Hospital.

All participants were given written information about the study. Participants in Study I and Study II signed consent forms ahead of data collection. All participants were informed of the aims of the study they were part of, of potential disadvantages in participating, and that they could withdraw from the sessions at any point, though no participants availed of this. Respondents in Study III could exit the online survey at any point.

4. Methods

This chapter begins with an overview of the methods applied in the thesis. Because the group interview data is essential for all three studies, I provide an outline of the sample, recruitment, and facilitation of the group interviews before moving onto the methodological aspects of each individual study. The three studies are presented consecutively, first looking at the methodological aspects of Study I, before providing a general overview of Q methodology, then presenting different aspects of Q methodology as they pertain to Study II and Study III.

4.1 A mixed-methods approach to eliciting public views

The aim of this thesis is to ascertain subjective views held by the public on the meaning of severity, and to study the prevalence and distribution of these views in the Norwegian population. There are many ways to elicit public views, both qualitatively and quantitively. The most widely used qualitative methods for public elicitation studies are one-to-one or group interviews, while quantitative methods can be classified along ranking, rating, or choice-based approaches (Ryan et al., 2001). To thoroughly explore public views on severity, this study is built on a mixed-methods research design and consists of both qualitative and quantitative components. This allows for both a broad, qualitative exploration of views held by members of the public, as well as a quantitative exploration of the distribution of those views, quantifying their prevalence across the Norwegian population.

Group interviews were conducted with a broad sample of members of the public, with the research team travelling across Norway to speak to individuals with different backgrounds and sociodemographic profiles. The group interviews were conducted in the format of open conversation, allowing for an exploration of both the breadth of views on severity and the depth of meaning within these views. The three aims of the thesis—examining views on severity, searching for potential clusters of opinions, and studying their distribution across the Norwegian population—were addressed in three steps: (i) a thematic analysis of the group interview data; (ii) applying Q methodology to look for clusters of opinions on severity, building on both group interview data and data from a rank ordering exercise with a separate set of participants: and (iii) applying Q methodology, via a Q-to-survey study, to identify the prevalence and distribution of these views.

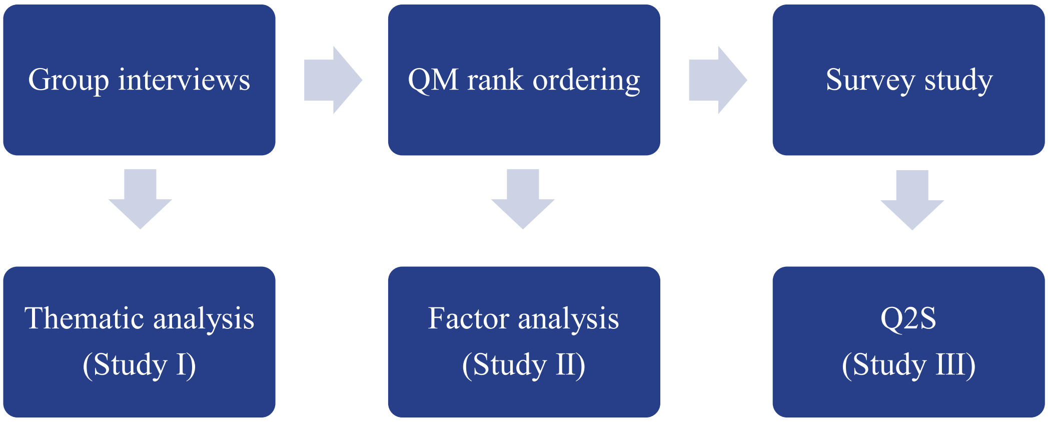
The thematic analysis forms the basis of Study I and is a qualitative exploration of the views on severity identified in the group interviews. I believe this is a well-suited approach to ensure an in-depth exploration of public views on a concept as complex as severity. This approach allowed for an exploration of what the participants were really talking about when asked about severity, and of the underlying views and conceptualisations expressed.

Another aim of SEVPRI was to explore whether—across all these views—there is any area of consensus, or any clusters of opinions that might represent different, possibly overlapping understandings of severity. As a study of subjectivity, Q methodology (QM) is well-suited to this aim. Combining qualitative and quantitative techniques, QM provides a method to systematically study subjective views on a given topic, and is intended for studies on subjective views, opinions, and beliefs (Watts & Stenner, 2022). The objective of a Q study is to identify shared viewpoints on a topic and to present those as different clusters of opinions on that topic (Brown, 1993). For a concept like severity, where there appears to be many different interpretations and no clear definition, QM is well-suited to allow for an in-depth, exploratory approach to identifying clusters of opinions.

While a Q study identifies and describes views on a topic, it does not say anything about the prevalence of those viewpoints. One can build on Q studies by conducting Q-based surveys, or Q-to-survey (referred in Q methodology literature as Q2S) studies. Q-to-survey studies build on the findings of a Q study, by extracting the central descriptions of a concept from the different clusters of opinions on it and implementing these into a survey (Baker et al., 2010). Using Q-to-survey methods and administering the survey across a representative sample of the population can provide knowledge on the prevalence and distribution of the viewpoints identified in the original Q study, across a population rather than only for the sample in the Q study. Q-to-survey methods enable verification and quantification of the views identified in a Q study, allowing researchers to draw some generalisable conclusions about public views on the concept being studied.

The studies within this thesis are presented in the following order: (i) thematic analysis, (ii) Q study, and (iii) Q-to-survey study. This is the clearest and most coherent presentation of the data and of the findings, as the thematic analysis provides a broad exploration of severity, the Q study explores clusters of opinions across different views on severity, and the Q-to-survey study presents results from a survey based on the identified clusters of opinions. The stages of analysis were, however, conducted in a different order. Following data collection in the group interviews, Study II was conducted. Following this, the survey for Study III was designed and administered. Thematic analysis for Study I was conducted after Study II was completed, and in parallel with the development of the survey for Study III. The rationale for this order was based on practical reasons, to ensure the PhD project could be completed within the set time frame. The Q-to-survey study for Study III is based on the results from Study II. Having completed Study II, there was ample time to return to the data from the group interviews and explore these in-depth for the thematic analysis in Study I. This was not problematic, as Study II and Study III were not designed to build on Study I. However, the results are presented in the order as outlined in Figure 1.

Figure 1 Overview of data collection and analysis.



4.1.1 Group interview data

The data from the interviews were studied separately for the thematic analysis (Study I) and for the Q study (Study II), and—as Q-to-survey studies build on results from Q studies—indirectly also for Study III. As these data were used across the studies, information about sampling, recruitment, and facilitation will be presented in the following section, before the three studies and their respective methodologies are presented individually.

Sample size and recruitment

In qualitative studies, there are no absolute rules for determining sample size, but rather depends on the line of inquiry and the purpose of the study (Patton, 2015). In the original design of SEVPRI, the anticipated number of participants necessary to reach saturation was expected to be approximately 60 individuals. This was based on established theory in Q methodology, where a sample size of 40 to 60 participants is generally considered to be appropriate (Brown, 1993) and on saturation according to data redundancy. Saturation determined by data redundancy entails establishing saturation once no new data emerges in the interviews, i.e., when the views of new participants appear repetitive of previously expressed views (Saunders et al., 2018). With a data-driven approach to saturation, data collection would conclude once comments and views began to repeat themselves, with no new views emerging (Francis et al., 2010). Saturation was considered throughout the data collection process and determined during data collection rather than during the subsequent analysis of the material. To ensure that a wide variety of views were identified, necessary both to conduct a meaningful thematic analysis and to ensure a Q study factor distribution that represents a genuine breadth of opinion, the aim was to speak to a diverse group of individuals. Sampling was conducted purposively, seeking to recruit participants across different demographics, with diverse backgrounds in relation to age, gender, education, socioeconomic background, and geographical location. Sampling purposively enabled a tracking of the backgrounds of participants, so that recruitment for future interviews could be directed towards demographic groups that appeared underrepresented in the sample. Participants were recruited across different locations in Norway and from both rural and urban areas, and we planned to conduct the group interviews across five different geographical locations (Oslo, Bergen, Trondheim, Tromsø, and Alta) (see Chapter 5.1 for further details on recruitment and data collection).

The study design of SEVPRI and this thesis was outlined in 2019, prior to the SARS-CoV-2 pandemic. Recruitment and data collection were conducted during the pandemic (February to July 2021). While the group interviews were originally designed as in-person sessions, alterations were made to accommodate pandemic restrictions. For this reason, both recruitment and data collection were conducted in different formats. Recruitment was done online via SEVPRI's social media accounts (Facebook and Twitter), sharing a link to an online recruitment platform, and by hanging posters advertising the study in shops and on lampposts around Oslo, as well as in waiting rooms of general practitioner offices in Bergen. Adapting to pandemic restrictions on physical gatherings, data collection commenced with online meetings via Zoom (Barbu, 2014). While the original design outlined six to eight participants per group, the decision was made to conduct the online groups with two to four participants. This was based on the nature of online meetings, in contrast to in-person ones: time lags, less non-verbal communication, and fewer opportunities for informal conversation allowing participants to become somewhat acquainted with each other, made the online group interviews more challenging. Conducting them with fewer participants helped to create a better environment for people to express their opinions and share their views and stories. Once pandemic restrictions were lifted, data collection was converted to in-person meetings. These were conducted with six to eight participants. Overall, whether online or in-person, conducting group interviews in the dynamic nature of a group setting allowed for stimulating discussion where participants could discuss ideas and engage with each other's views (Malterud, 2012).

All participants received a universal gift card as compensation for the time they spent in the group interview. Online participants, who were not required to travel, received a gift card of NOK 250 (approximately €23). In-person participants received NOK 500 (approximately €45), as their participation required travel to the meeting locale and demanded more of their time. The online sessions were expected to last one to two hours, while the in-person sessions, with more participants and more logistical challenges, were expected to last two to three hours. The recruitment period for all group interviews was from February to July 2021.

Conducting the group interviews

The aim of the group interviews was to capture the breadth of opinions on severity, ensuring as many different views as possible were gathered and explored. For this reason, the group interviews were not moderated to reach consensus, but to explore the different views participants held. We wished to avoid influencing their answers as much as possible, and to create an atmosphere where participants felt comfortable enough to be forthcoming with their opinions and to engage in discussion with each other when they disagreed. Ahead of data collection, a group interview guide was created (see Appendix A). This included a brief introductory text to the topic of severity, which every group interview was opened with. Some necessary context was provided to explain why knowledge on severity is important, and how severity is relevant to priority setting. To avoid influencing participants' views, the introductory text was kept to a minimum, without any statements surrounding its potential meaning. A pilot interview was conducted with members of a user panel at Akershus University Hospital, which is a group of individuals consisting of health personnel, patients, and members of the public. The interview guide was found to be sufficient, with an understandable introduction and instruction for the conversation.

The group interviews were facilitated mainly by me, or by another member of the SEVPRI research team (either the WP3 PhD candidate or the SEVPRI Principal Investigator). The facilitator would introduce and lead the group interviews, asking follow-up questions and bringing up new topics from the interview guide. The other team members functioned as observers but could also ask follow-up questions and pick up on new themes the facilitator might have missed. Following the introductory text, and after collecting oral and written consent, the group interviews were started with an open question from the facilitator. The facilitator would ask the group what severity means to them, and what their first thoughts were. The opening question was open-ended to encourage participants to freely share their first views on and associations with severity, and to identify perspectives participants felt were relevant (Alvesson & Sköldberg, 2017). Each participant was given the opportunity to answer the opening question. The group interviews were intended to be conducted in the format of open conversation—and will henceforth be referred to as such—guided by the participants themselves and their discussions with each other. The conversations were guided by the issues participants raised and what they felt was relevant to the meaning of severity. The role of the facilitator was to ask clarifying questions, to get a deep understanding of the views the participants expressed, as well as to ask participants about some attributes of severity considered to be relevant. To ensure this was done in a consistent and transparent manner, the open conversation style was supported by a topic guide, also part of the interview guide, intended to ensure a breadth of topics were covered during the conversations. The topic guide was created ahead of data collection, following a comprehensive literature search on the subject of severity (Barra et al., 2020). The topic guide was thought to cover potential attributes of severity considered relevant following the literature search, such as age, death, pain, stigma, etc. The facilitators from the SEVPRI research team sought to allow discussions to develop organically, and explored the topics participants brought up, but also asked questions regarding the list of attributes in the guide to gauge participants views on these. The topic guide was dynamically updated to include emerging themes participants brought up in the conversations. The topic guide was not made available to participants, but rather intended to guide facilitators.

At the end of the conversations, participants were asked to fill out a questionnaire (see Appendix B) about their socioeconomic status, self-reported health status, situations that have affected their views on severity, as well as feedback on the conversations. In addition, participants were asked to fill out a descriptive survey of self-reported health, using the EQ-5D-5L instrument, developed by the EuroQol Group (EuroQol Research Foundation, 2023). In the EQ-5D-5L, respondents describe their health according to the five dimensions of mobility, self-care, ability to participate in usual activities, degree of pain/discomfort, and degree of anxiety/depression. In the in-person conversations, the questionnaires were filled out by the participants. In the online conversations, participants were called by the facilitator via telephone following the conclusion of the conversations, and the facilitator would read out the questions in the questionnaire and fill it out on their behalf. All the conversations were digitally audio recorded. Following conclusion of data collection, these were transcribed *ad verbatim* in Norwegian. I transcribed fourteen of the conversations, while seven were transcribed by other members of the SEVPRI team.

4.2 Thematic analysis (Study I)

The analysis in Study I is the result of systematic text condensation. This is a qualitative approach to text analysis where the data are subjected to repeated cycles of induction and deduction (Malterud, 2019). As with other forms of thematic analysis, the aim is to search across a data set to identify patterns of meaning (Braun & Clarke, 2006). This analysis is data-driven with a strong inductive element, meaning the identified themes are strongly linked to the data themselves rather than a specific (theoretical) framework. This form of analysis tends to provide rich descriptions of data (Crabtree, 1999) and allows for the identification of both semantic and latent themes (Braun & Clarke, 2006). With the semantic approach, themes were identified based on what participants said and discussed, and the topics they brought up. At the same time, it was important that latent themes could emerge, i.e., underlying ideas and conceptualisations within the issues participants discussed, which were perhaps not labelled explicitly but which seemed important in light of the underlying meaning being conveyed.

To ensure congruity in the analytical process, three of the four co-authors for Study I (me, JR, and HL) each coded three of the transcripts, independently. We subsequently compared the coding we had performed. The analysis itself was conducted using NVivo (release 1.6.1). I conducted the four stages of the analysis (see Table 2), with all co-authors contributing to discussions along the way.

Table 2 The four stages of analysis for conducting the thematic analysis (table adapted from Paper I).

Stages	Description of analytical process
1: From chaos to codes: read-through	Getting familiar with the data by reading through all transcripts. Note-taking using mind maps to record topics for potential codes.
2: Coding the material: deductive-inductive cycles	Coding all transcripts, adapting the codebook as necessary. Dynamically developing the codebook during the coding process (inductive).
3: From code to meaning: identifying themes	Studying the codes in isolation and in conjunction with each other, searching for themes. Creating mind maps of potential themes and identifying if and how codes fit within these.
4: From de-contextualisation to recontextualisation: descriptions	Connecting the themes to broader body of literature, looking for connections within and between themes. Recontextualising by returning to transcriptions to consider if themes reflect what participants discussed. Writing out narrative within themes.

Stage 1 of the analysis involved a preliminary read-through of all the transcripts. Some preliminary notes were taken during the reading, identifying topics and placing these onto mind maps (Sullivan & Forrester, 2018). In the Stage 2, the transcripts were re-read, while this time also being coded. The codebook consisted of the topics identified in Stage 1, applied deductively during coding, as well as new codes that appeared inductively during the coding process. The codebook was thus updated dynamically throughout Stage 2. Stage 3 involved a new read-through of all the transcripts, once again to ensure all relevant data were coded. Following this, the transcripts were studied inductively to identify overarching categories. In Stage 4 of the analysis, the categories identified in Stage 3 were connected to relevant bodies of literature, to identify meaning within and across categories.

Disengaging from the initial analysis

Though Study I is presented first in the thesis, as mentioned above, the analysis for Study II was conducted first. I had already analysed the data from the conversations for the purpose of Study II when commencing the analysis for Study I. To provide a thorough thematic analysis and an accurate presentation of the data for Study I, I had to separate myself from the analysis I had previously undertaken. This required a rigorous and conscious effort.

The first analysis of the conversation data, which is explained in Chapter 4.3.1, involved a thorough search through the transcripts for statements relating directly to severity. This was not an analysis of the views and opinions expressed in the conversations, but a search for raw statements specifically pertaining to severity. This stage did not involve active engagement with the content, the views expressed by participants, or with the overall meaning(s) of severity that emerged. This was intentional, as the purpose of this exercise was not—unlike in a thematic analysis—to uncover the meaning behind what had been said by participants. Preliminary stages of a Q study analysis allow for less engaged involvement with the data, with a more mechanical 'search engine' approach.

Once the analysis and broader work on Study II was completed, I returned to the transcripts to conduct the thematic analysis for Study I. Reengaging with the transcripts for the purpose of this analysis commenced approximately one year after the initial Q study analysis, allowing me to return with relatively fresh eyes and consider the data anew. Furthermore, the actual data collection and transcription phase was by far the most time-consuming exercise of the PhD project, meaning that the time I originally spent with the data was considerably longer than the time spent conducting the analysis for Study II. Consequently, I felt I was more engaged with the data collection process than I was with the previous analysis—because I had spent so much time with data collection and transcription. Nonetheless, I was wary of bringing the results of the Q study analysis into the thematic analysis for Study I. I worked actively to disengage from this, reading through the manuscripts many times to reengage with what participants had shared, rather than the results from the Q study analysis, and—for the first time—to consider what the overall meaning and potential themes across participants' views might be. It was also helpful that the co-authors on Study I had not participated in the group interviews or read the transcripts and were unfamiliar with this material (DGTW is a co-author on the paper from Study II, but did not participate in data collection or transcription). Their questions and input regarding the material and analysis was helpful in separating the analyses for the two studies.

4.3 Q methodology

Q methodology is a mixed-methods approach to studying subjective viewpoints on a given subject. It was pioneered by William Stephenson, who sought to develop a method for the systematic study of subjectivity (Stephenson, 1935). It provides techniques for the exploration of an individual's opinions and beliefs (Brown, 1993). By asking a group of individuals to rank order a set of statements expressing views on a subject, and subsequently performing factor analysis on these rankings, the resulting factors are thought to represent clusters of opinions—or subjectivities—on a topic. A Q study thus describes a population of viewpoints, rather than a population of people (Risdon et al., 2003); Q methodology employs individuals, rather than variables, as tests (Brown, 1993). Accordingly, the methodology is also referred to as inverted, or by-person, factor analysis. This approach enables the researcher to not only explore viewpoints on a topic, but to identify and describe shared viewpoints between people (Watts & Stenner, 2022). By identifying clusters of opinions on a topic, one can describe areas of consensus across the material, as well as identify the differences between these clusters. Q methodology thereby provides techniques to gather detailed information about shared viewpoints across a topic, and an overview of existing views.

4.3.1 Conducting a Q study

Q studies generally consist of five steps: (i) defining a concourse, (ii) developing a Q set, (iii) sampling participants (the P set), (iv) performing a Q sorting exercise, and (v) analysis and interpretation of the Q sorts (van Exel & de Graaf, 2005). In the following section, the steps involved in a Q study are described in this order. Common terms used in Q methodology are presented in Table 3.

Table 3 Summary of common terms used in Q methodology.

Terms	**Explanation**
CoI	Condition of instruction (Q sort guide)
Concourse	The discourse on the topic in a Q study (collected from the public discourse, the literature, music, art, etc.); defined according to what the topic is
P set	Full group of participants in the Q sort exercises
Q deck	Full set of statements for the Q sort, each represented on a separate card totaling in a deck
Q set	Full set of statements for the Q sort
Q sort	Exercise of ranking statement cards and placing them onto a grid
Q2S	Q-to-survey study (development of a Q study into a survey)
Z-score	Score given to each statement in a factor, representing the weighted average score that statement was given by participants who were defining of that factor

The concourse and the Q set

The statements that participants eventually rank order during the Q sorting exercise are collected from what is referred to as the concourse, or the ‘flow of communicability’ surrounding a topic (Brown, 1993). The concourse can be statements of opinions, but also objects, pictures, and other forms of discourse (McKeown & Thomas, 2013). The Q study conducted for this PhD project is centred on a verbal discourse, thought to represent a discourse of accounts of severity. Verbal discourses can typically be explored through interviews, newspapers, books, etc. The concourse provides the raw material for the Q study. To gain an insight into the concourse via interviews, one ought to speak to individuals thought to have something to say about the issue in question (Watts & Stenner, 2022).

In a study of public opinion, exploring the concourse involves interviewing members of the public. The role of the researcher is to define the relevant concourse for the topic of their study, and to extract statements from this concourse that accurately represent the breadth of expressed opinions on the topic, referred to as a Q set (van Exel & de Graaf, 2005). The Q set should be representative of the overall concourse, and the researcher thus aims to select statements which are distinct from each other and represents the full spectrum of different opinions within the concourse (Baker et al., 2017). The selection of the statements for the Q set is done at the discretion of the researcher, and can be done both inductively (selecting statement as you go through the concourse) or deductively (basing selection on a pre-determined theory). Studies comparing different approaches to selecting statements for the Q set suggest that they converge on similar conclusions (Thomas & Baas, 1993). A Q set typically consists of 40 to 80 statements (Stainton Rogers, 1995), which are printed on individual cards. These cards are referred to as the Q deck, which is used in the Q sorting exercise.

The P set and Q sorting

The group of participants who perform the Q sort exercise are referred to as the study’s P set (McKeown & Thomas, 2013). The P set consists of individuals expected to have opinions on the subject matter, and generally consists of fewer individuals than there are statements (van Exel & de Graaf, 2005). At the outset of the Q sort exercise, participants are introduced to the topic at hand through an introductory text, with instructions on how to perform the task. This is referred to as the condition of instruction (Damar & Sali, 2022). Participants are provided with the Q deck, with each statement represented on an individual card, as well as with a grid on which to place the cards. The grid can have different forms, but commonly has a quasi-normal distribution. The grid, also referred to as a score sheet, represents a continuum of agreement, ranging from, for example, ‘most agree’ to ‘most disagree’. The kurtosis of the distribution depends on the subject at hand, and will be steeper for topics expected to involve a lot of ambiguity, or where few of the statements are considered particularly salient (van Exel & de Graaf, 2005). This gives participants more space to place cards they are less certain about around the middle, than at the two extremes. The distribution usually ranges from around -5 to +5 (Brown, 1993), but will depend on the size of the Q deck.

Participants are instructed to place the cards onto the grid, usually with one space for each card through a forced distribution. Both forced and free distribution can be employed, with comparisons of the two suggesting a negligible effect on results (Brown, 1993; Hess & Hink, 1959). Via the condition of instruction, participants are asked to first read each statement, and to divide these into three piles: statements they generally agree with, statements they generally disagree with, and statements they feel neutral about. Following this, they are asked to consider the piles one by one, and place all the cards from one pile onto the grid before addressing the next pile. Participants can make changes to their card placements throughout the Q sorting exercise (Watts & Stenner, 2022). Once all statement cards in the Q deck are placed onto the grid, the participants are usually interviewed, either verbally or via questionnaires. They are asked to elaborate on their opinions on the most salient statements, and why they distributed the statements in the way they did. This final step provides important information for the subsequent analysis and a further layer of detail to understanding participants’ views.

Factor analysis and interpretation

Once the Q sorts are collected, factor analysis can commence. The ideal factor solution is determined according to two elements: the statistical qualities of the solution, and the qualitative interpretability of the factors (McKeown & Thomas, 2013). Analysis is generally conducted using dedicated software, where the statistical analysis is based on the calculated correlation matrix of all the completed Q sorts, representing the degree of (dis)similarities between the participants in the Q sort exercise (van Exel & de Graaf, 2005). This exploratory factor analysis identifies different groupings of Q sorts based on the similarities and differences between them all. This is an analysis of how many different Q sorts there actually are, and the degree to which there is overlap between them (Brown, 1993). The factors are subjected to rotation, examining the factors from different angles and aiming for a final factor set in which each factor represents Q sorts with a high level of correlation, and which is distinct from other factors (Exel & de Graaf, 2005). With exploratory factor analysis, different factor solutions are explored, based on the number of factors one asks the software used to conduct the analysis to provide. Deciding on a final factor solution involves considering the number of flagged Q sorts within each factor—the flagged Q sorts representing the sorts significantly associated with each factor—and comparing the different factor solutions and whether new factors emerge as the number of factors is increased (Watts & Stenner, 2022).

Every statement in the Q deck receives a factor score (referred to as Z score), representing the weighted average score given to that statement by participants deemed defining of that factor (Watts & Stenner, 2022). Based on Z scores, statements are placed onto the grid which now represents the composite sort for that factor. Within a *k*-factor solution, the *k*-factor composite sorts thus represent a weighted compromise of the flagged Q sorts on each factor. Each Q sort from the Q sort exercises is compared to the composite scores, and based on the rotation performed and the amount of factors selected, each Q sort has a loading on the different factors. The statements within the Q sorts are also analysed according to whether or not they can be considered distinguishing statements (ranked significantly higher or lower on one factor than on other factors) or consensus statements (not distinguishing between any of the factors).

The next analytical step involves a qualitative analysis of the composite sort of each factor, and interpretation of the statement placement within it. Special attention is given to distinguishing statements and to salient statements (those placed at the two extremes of the grid). The consensus statements are used to describe potential similarities between the factors. Finally, the post-Q sort interviews conducted with participants are analysed, providing detailed information helpful to the interpretation of the factors. These supplementary data are employed to assess the relevance and feasibility of different factor structures, and enable richer description of the factors (Baker, et al., 2010).

Q studies result in rich, detailed descriptions of a smaller number of shared viewpoints, or clusters of opinions, represented by the different factors and their composite Q sorts.

Q-to-survey studies

Most Q studies typically conclude with the description of the factors in the final factor structure. Although these studies present detailed information about views of the public on a given topic, they do not provide information about the viewpoints' prevalence in, or their distribution across, the general population. Conducting Q sort exercises is time consuming and difficult to perform with a representative sample of a population. For studies addressing questions of prevalence and distribution, survey instruments which can be administered more efficiently are desirable (Baker et al., 2010).

Capturing social views and values has its distinct challenges, and there is an increasing body of research on effective methods to solicit and measure public views (Schoon & Chi, 2022). Because Q studies are particularly suitable to probe subjective accounts of views and values, but less suitable for surveying purposes, there is an active field of research involving development of Q based surveys methods. These survey methods provide data on the representation and distribution of viewpoints identified in a Q study, and enable analysis of the association between the different viewpoints on a topic and sociodemographic characteristics. Designing surveys based on Q studies enables the researcher to build on the rich, detailed descriptions of different viewpoints represented by the factors generated in a Q study. By extracting the defining features of the different factors, these can be represented in a survey study to capture the subjective views of the broader population on a given topic.

There are several approaches to developing surveys based on Q studies. These revolve around three different methods: Talbott's Q block, Brown's standardised factor index score, and self-categorisation to abbreviated factor descriptions (Baker et al., 2010). Talbott's Q block involves asking participants to rank order a selection of statements from the original Q study, selected based on their salience and distinctiveness (Talbott, 2010). In Brown's standardised factor index score, participants are asked to score a set of statements representing different factors, which are subsequently used to compute standardised index scores for each factor (Baker et al., 2010; Brown, 2002). While both of these methods have been applied (Baker et al., 2014; van Exel et al., 2006), they also have an important limitation: selecting only a few statements to represent an entire factor raises questions of how representative these are of the factors identified in the original Q study. To address this concern, a third method of self-categorisation has been developed in recent years. Here, participants are presented with a summary description of each of the factors in the original Q study, presented in text-form. Participants are asked to indicate the degree to which each of the descriptions align with their views. The self-categorisation approach is intended to provide a more wholesome representation of the factors. This approach is not limited to only a few statements, and does not require that participants consider a set of statements outside their factor context (Baker, et al., 2010). The self-categorisation approach is still in development, with several studies contributing to the body of research employing self-categorisation (Donaldson et al., 2010; Jedeloo et al., 2010; van Exel et al., 2011).

Which method one selects to conduct a Q2S study ought to be considered in light of the question the researcher seeks to answer. If the question is how a larger group of respondents would sort a set of statements, without having to conduct the time-consuming full Q sort exercises with a large group, both Talbott's Q block and Brown's standardised factor index score are well-suited. If, however, the aim is to consider the preferences of a larger group of respondents when faced with descriptions of factors from a Q sort study, self-categorisation is arguably preferable. Rather than presenting respondents with excerpts from original Q studies with a select group of statement, self-categorisation studies provide condensed, coherent representations of the factors, seeking to communicate the fuller meaning each factor is thought to represent. In the survey study presented in this thesis, the latter of the three methods is applied.

4.3.2 Identifying accounts of severity (Study II)

Study II is an in-depth empirical investigation of subjective views on severity. Q methodology was applied to study the viewpoints of participants, elicited via a three-stage process involving the steps of Q methodology as outlined in Chapter 4.3.1 (Brown, 1993).

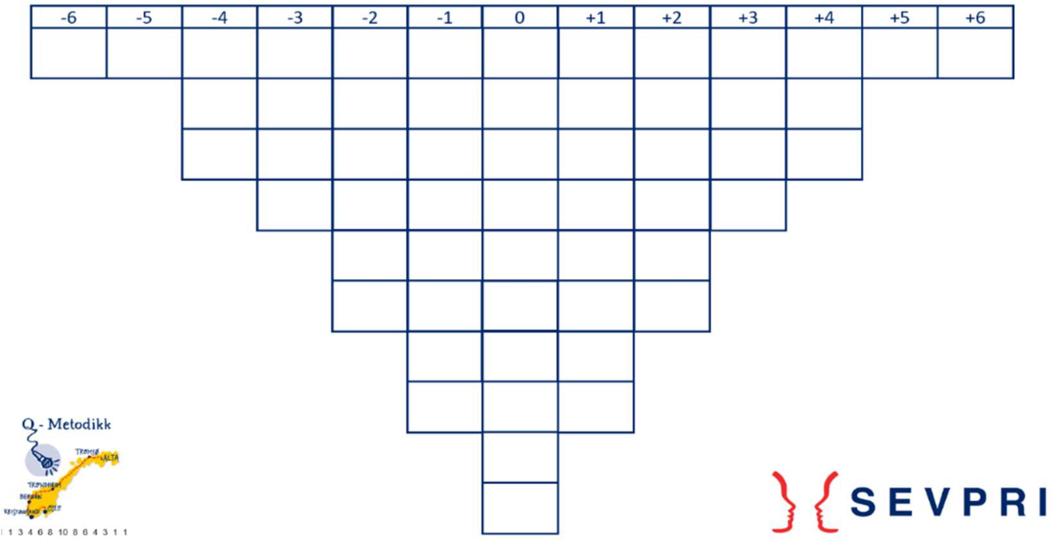
Stage 1: Deriving the Q set

The statements for the Q set were collected from the transcripts of the conversations, by a process of coding the transcripts and collating all statements that expressed an opinion on severity (Baker et al., 2017). This was followed by an analytical process involving all those statements, now out of the context of the conversations but looking to identify anything that was said and could be interpreted as a proposition about what 'severity' means. All statements pertaining to severity were collected and categorised according to the themes in the topic guide, which had been dynamically updated throughout the data collection process as participants brought up new topics. To condense these into a final Q set, we eliminated duplicate statements and statements that would be difficult to understand out of their conversational context. The overall aim of this exercise was to condense the data from the conversations into representative statements, preserving coverage of all the topics discussed in the conversations. A coding procedure, giving every conversation and participant a unique code, was used to ensure we did not select statements from only some of the conversations or participants.

Stage 2: Q sort exercises

Participants in the conversations provided raw data for deriving the Q set. A new set of participants (with no overlap from those in the conversations) provided raw data for the following factor analysis, based on these participants' Q sorts. In the Q sort exercise, participants were provided with the statements in the Q set as well as a Q sort grid (see Figure 2). With a fixed grid such as this, there is a forced distribution of one card per square, as is common to Q methodology (Watts & Stenner, 2022).

Figure 2 The grid used in the Q sorting exercise (figure from Paper II).



The Q sort exercise was piloted with colleagues to ensure the instructions were understandable and the task manageable, and to explore if there were any views considered to be missing from the statement set. The Q sort exercises were conducted in a group setting for the sake of efficiency, but each participant completed the exercise undisturbed, individually, and with the facilitator available to answer questions.

The sessions were opened with a brief introduction, similar to the introductory text in the conversations. Participants were then asked to sort all their cards into three separate piles (‘agree’, ‘disagree’, and ‘neutral’), before they were taken through the instructions of placing all the cards onto the grid, one pile at a time (see Appendix C for introductory text and instructions). Following the completion of the Q sort, participants were asked to elaborate on their rankings and viewpoints in a separate questionnaire (see Appendix D), as well as to complete the same demographics questionnaire as distributed in the conversations (see Appendix B).

Stage 3: Factor analysis and Q sort interpretation

The first step of a Q analysis is mainly statistical, but also has a qualitative element in the exploration and selection of a final factor solution (i.e., number of factors), as outlined in Chapter 4.3.1. Analysis was conducted using the Q methodology software package KenQ Analysis Desktop Edition (KADE) version 1.2.1 (Banasick, 2019), with centroid factor analysis with varimax rotation. We applied conventional Q methodology practice to determine the factor solution. We compared the different factor solutions, examining whether new factors emerged as the number of factors was increased (Watts & Stenner, 2022). This process was somewhat abductive in its nature (Douven, 2021). The aim was to find the best possible explanation for the selected factor solution, i.e., a solution in which each factor, and how the statements were placed within each factor’s composite sort, provides the best explanation of the viewpoint they represent. This process was conducted via deliberation between all co-authors.

Deciding on a factor solution also entailed consideration of the number of flagged Q sorts within each factor, seeking to have at least four flagged Q sorts for each factor. The degree to which participants, via their Q sorts, loaded on a factor was determined by whether they (i) had a statistically significant factor loading, and (ii) that more than 50% of the communality corresponded to that factor. The factor loadings thus determined flagged sorts (i.e., which participants’ Q sorts were used to generate the composite sort of each factor), and the weights attached to the flagged respondents’ sorts. A flagged respondent is someone whose Q sort is mostly (>50%) explained by one sole factor. See attached paper for fuller description and tables with statement scores and factor loadings.

Once a factor solution had been decided on, qualitative analysis of the factors based on the composite Q sorts could begin. After analysing and discussing the different factor within the study team, a final interpretation of each factor was brought together. Analysis of the composite Q sorts involved consideration of the placement of all statements, with particular emphasis on the salient statements, distinguishing statements, and consensus statements. I also analysed the information provided by participants who were flagged on the factor from the post-sort questionnaires, to ensure a rich understanding and solid interpretation of their views.

4.3.3 Studying the distribution of viewpoints on severity (Study III)

Design

The viewpoints (i.e., the factors) identified in Study II were converted into short vignettes to be presented in the survey study (Study III). These were written to resemble the factors as closely as possible, with emphasis on the salient and distinctive characteristics of each viewpoint (Mason et al., 2016). This entailed a consideration both of those features in the viewpoints that the participants in the Q study felt strongly about, and distinguishing features between the viewpoints. The vignettes also had to be suitable for presentation in a survey, with an aim of approximately 100 words per vignette (Hughes & Huby, 2004). The four vignettes are presented in Table 4.

Table 4 Summary descriptions of the four viewpoints from Study II, referred to as 'vignettes' (table adapted from Paper III).

Vignette name	Vignette text
I: Lifespan	I think severity is about how health problems affect the natural course of life, and affect the natural development of life—and especially if it affects the young. Death isn't necessarily severe, especially when you're old, and there are things that are more severe than death. Pain isn't necessarily severe either. But things like mental illness and loss of dignity, that is severe. Severity is first and foremost when illness affects the natural course of life's different phases, and takes one's possibility to experience what you should be allowed to expect from life.
II: Subjective	I think severity is almost entirely about how one experiences health problems, and severity can't be tied to any specific diagnoses or conditions. Severity depends completely on the experienced situation, and what one feels is severe for oneself and one's life. You can't define severity objectively. There is no right answer to what severity is, but when you are affected at a young age, or by something that just gets worse and worse, or leads to a loss of dignity, that makes it more severe. And it's maybe more severe if it affects a parent who is responsible for a child. But it's hard to say anything definitive about what severity is.
III: Objective	I think severity has to be defined by some objective measures, like age, diagnosis, prognosis, and urgency. It can't be up to each individual to decide what is severe and what isn't. We need some criteria, and health personnel or other experts should be involved in making those criteria. Both mental illness and pain can be severe, or not being able to work, but severity is not about what someone believes, feels, or thinks. The degree of severity is decided by objective facts about the condition or diagnosis that one has.
IV: Functioning and Quality of Life	I think severity depends on how it affects your day-to-day life: if it alters your functioning at work, at home, your ability to participate in society, enjoy your hobbies and things like that. Still, not just anything can be severe. There are clearly some objective links between a health problem (mental or physical) and how your quality of life is affected. But severity has to be measured by how it affects your quality of life and your levels of functioning.

Notes: The English translation is for this exposition only. The items were presented to Norwegian-speaking respondents in their original language.

The survey opened with an introductory text similar to the text used in the conversations and Q sort exercises. Participants where then asked to perform a ranking exercise with different statements from the Q deck. That exercise is not part of this thesis, but was used as a warm-up exercise and may be used for the development of a new Q-to-survey method in future work. The main section of the survey contained one, two-part task: a vignette ranking task and a vignette scoring task. In the vignette *ranking* task, the four vignettes were first presented individually, in randomised order, before they were all presented on the same page (in the same, randomised order). Respondents were asked to rank the four vignettes from "most like my view" to "most unlike my view". In the vignette *scoring* task, respondents were presented with the four vignettes in the order they had provided in the ranking task. They were then asked to score the vignettes on a visual analogue scale (VAS), anchored at "completely descriptive of my view" at the top to "completely different from my view" at the bottom. The VAS was not numbered but contained nine evenly spaced marks[1]. Finally, respondents were asked to describe, in their own words, their views on severity, as well as to provide feedback on the survey[2].

Data collection

The vignettes were first piloted with SEVPRI's Advisory Board, for feedback regarding wording and intelligibility. A preliminary version of the survey was piloted with the user panel at Akershus University Hospital, to receive feedback on the clarity of the tasks and wording. The survey was then distributed online via the market research company Norstat (www.norstat.no), running two pilots with approximately 100 respondents each ahead of the final survey distribution.

Norstat uses quota sampling via their panel, consisting of Norwegian adults from 18 to 99 years old, delivering target quotas for representation in terms of age, sex, and region. Norstat also provided additional information about respondents regarding gross household income and a centrality index, the latter of which describes whether respondents live in rural or urban areas. We sought to collect approximately 1000 responses, in addition to the pilot data. All respondents in the final survey were entered into a lottery that had two prizes, where each prize was a NOK 5000 (approximately €450) gift card.

Analysis

The vignette ranking and scoring tasks were treated as a single, two-part task, where the VAS scores from the scoring task were used to define alignment with the vignettes. Vignette

---

[1] The VAS shown to respondents contained 9 evenly spaced marks. In the data provided by Norstat, the scale was converted to an 11-point scale (0–10), with the VAS scores spaced out in the same manner but according to 11 points rather than 9.

[2] For full survey (available in Norwegian), visit: https://web.norstatsurveys.com/survey/selfserve/53c/2308137

alignment represent respondents' endorsement of the vignettes. Respondents' strength of alignment with the vignettes was considered according to four different approaches, or categorisations of analysis. The first analysis placed all respondents onto a vignette based on the highest VAS score. If a respondent tied two or more of the vignettes, they were assigned to all their top vignettes. The second analysis, thought to represent stronger vignette alignment, demanded that respondents expressed a certain level of support for their highest scored vignette. In this analysis, respondents were placed on vignettes based on the requirement that the vignette(s) was scored ≥7 on the VAS. If more than one vignette was scored ≥7 on the VAS, they were assigned to all the vignettes scored ≥7. The third analysis, representing an alternative approach to stronger vignette alignment, did not require scores above a certain point on the VAS, but placed respondents on only one vignette. In this analysis, ties in the scoring subtask were resolved by referring back to the ranking subtask (where respondents could not tie vignettes, but had to rank order them to proceed). The fourth analysis, representing the strictest criteria and what we defined as vignette membership, required at least a score of ≥7 on the VAS, disallowed ties (i.e., respondents who tied vignettes ≥7 were excluded from the analysis), and if more than one vignette was scored ≥7, a gap requirement was introduced. This entailed that, to assign vignette membership, the top vignette had to be scored at least two points higher on the VAS than the next vignette.

To evaluate how representative the sample was, the sample and the target quotas for representation delivered by Norstat were compared to open source data from Statistics Norway (Høydahl, 2020). Finally, we fitted four linear regression models, with one for each of the vignettes, based on respondents' VAS scores and their sociodemographic characteristics. The aim of this analysis was to explore potential connections between vignette alignment and socioeconomic profiles. All statistics were computed with RStudio (RStudio Team, 2020).

5. Results

This chapter provides a summary of the key findings from each study. For a detailed presentation of results, please see the attached papers.

5.1 What are we talking about when we talk about severity?

The aim of Study I was to explore what members of the public talk about when discussing severity. The paper is titled *'"It's hard to say anything definitive about what severity really is"*: lay conceptualisations of severity in a healthcare context'. The recruitment period for the conversations that provided data for Study I was February to July 2021. The first 13 conversations were conducted online via Zoom (Barbu, 2014) (February to March 2021), and the final seven conversations were conducted in-person once restrictions were lifted (May to July 2021). The conversations were conducted at five different geographical locations (Oslo, Bergen, Trondheim, Tromsø, and Alta). Data saturation was established after 21 conversations, with a total of 59 adult participants (see Table 5).

Table 5 Participant demographics for the conversations (table adapted from Paper I). Values are numbers (percentages).

Characteristic	Participants (n=59)
Age category (years)[a]	
18-30	9 (15)
31-50	13 (22)
51-66	24 (41)
67+	11 (19)
No response	2 (3)
Gender	
Female	38 (64)
Male	19 (32)
Other/prefer not to say	2 (3)
Do you consider yourself religious or spiritual?	
Religious and/or spiritual: active in a congregation	11 (19)
Religious and/or spiritual: not active in a congregation	14 (24)
Neither religious nor spiritual	33 (56)
No response	1 (2)
What is your highest completed education level?	
Elementary/Upper secondary (up to 19 years of age)	9 (15)
Undergraduate degree/Apprenticeship	21 (36)
Graduate degree/PhD	27 (46)
No response	2 (2)
Have you or anyone you know well had severe illness?[b]	
Transient	18 (31)
Chronic	30 (51)
Deadly outcome	42 (71)
No response	0 (0)
How do you view your own health?	
Very good/Good	37 (63)
Just fine	15 (25)
Bad/Very bad	6 (10)
No response	1 (2)

[a] Age was given in one of the listed age-brackets.

[b] Categories are not mutually exclusive.

There was an overrepresentation of women in the sample. However, the sample was considered satisfactory considering that the aim was to capture breadth regarding sociodemographic background, with a qualitative exploration of views on severity (i.e., the aim was not to have a representative sample of the Norwegian population). Participants appeared open and enthusiastic to sharing their views on severity, which led to lively discussions. They evidently had a lot to say about severity and were forthcoming about finding it a complex issue. Participants were reluctant to categorise any conditions or situations as non-severe. Following a thematic analysis according to the four analytic stages outlined in Chapter 4.2, three themes were identified. These were interpreted to represent three different conceptualisations of severity.

Severity as subjective experience

In this conceptualisation, severity was seen as relating to the individual's own experience of their situation. Severity was perceived as inherently subjective, and only the individual suffering from a condition could know the severity of that condition, meaning there could be no general definition of severity. It was also expressed as inappropriate for severity to be determined by outsiders, including healthcare professionals and policymakers, as they do not have the personal experience of illness that a patient has. We observed that severity was associated with fairness, and that severity ought to be determined in a way that was fair to the individual. It appeared to be unfair to determine severity based on outsider views, and that it is the right of the individual patient to decide what is severe for them.

Severity as objective fact

The second conceptualisation contrasts the former and was centred on severity as objective fact. This conceptualisation involved an extrapersonal position, conceptualising severity beyond individual experience. The notion of criteria was central, i.e., severity was expressed as an objective notion determined by certain objective criteria—though participants disagreed on what these criteria might be. Severity was expressed as relating to something concrete and definable, with individual evaluations considered secondary to the objective severity of a condition. Fairness was again central but, in this conceptualisation, it took a different form. The emphasis was on ensuring fairness across all patients, which demanded the use of objective criteria rather than depending on the individual's own perceptions and views.

Severity as situation dependent

In the final conceptualisation, emphasis was on the context surrounding the patient. Three different subthemes were identified. In the first, severity was conceptualised via the social effects of illness on the patient, such as their ability to work and continue with hobbies. An illness was perceived as severe depending on how it affects a patient's relational, social, and work-related circumstances. In a second subtheme, severity was conceptualised by how illness affects those surrounding the patient, considering effects on next-of-kin, family, and friends. Parenthood was perceived as particularly relevant, with illness considered more severe if it impacts the dependents of patients. In a third subtheme, severity was conceptualised by the wider effects of illness on society. Here, issues like high treatment costs, productivity loss, and opportunity cost of care were relevant, associating severity with costs outside the healthcare sector. In this third subtheme, fairness appeared again—this time as a concern for determining severity fairly for society, considering what is fair at a societal level.

5.2 Clusters of opinions on severity

The aim of Study II was to identify different clusters of opinions on severity, representing distinct viewpoints. The title of the corresponding paper is, 'A severely fragmented concept: Uncovering citizens' subjective accounts of severity of illness' (Stenmarck et al., 2023). A Q methodological analysis of the transcripts from the conversations provided a total of 450 statements directly concerning severity. These were examined for intelligibility and relevance and distilled to a final sample of 53 statements representing the views that had been expressed in the conversations. Three further statements were added by the research team, which were conceived to represent views that were considered to have theoretical relevance, based on a comprehensive review of the literature (Barra et al., 2020), but which had not been expressed by participants in the group interviews.

Following a pilot of the final Q deck of 56 statements, instructions were edited slightly to make the exercise more understandable for participants. Recruitment and data collection (n=34) for the Q sort exercises were completed between January and March 2022 in Oslo and Bergen (see Table 6). Again, there was an overrepresentation of women in the sample, but here, too, the aim was breadth in sociodemographic background over representation.

Table 6 Participant demographics for the Q sort exercises (table adapted from Paper II). Values are numbers (percentages).

Characteristic	Participants (n=59)
Age category (years)[a]	
18-30	16 (15)
31-50	9 (22)
51-66	7 (41)
67+	2 (19)
No response	0
Gender	
Female	24 (71)
Male	9 (26)
Other/prefer not to say	1 (3)
Do you consider yourself religious or spiritual?	
Religious and/or spiritual: active in a congregation	2 (6)
Religious and/or spiritual: not active in a congregation	4 (12)
Neither religious nor spiritual	25 (74)
No response	3 (9)
What is your highest completed education level?	
Elementary/Upper secondary (up to 19 years of age)	8 (24)
Undergraduate degree/Apprenticeship	16 (47)
Graduate degree/PhD	10 (29)
No response	0
Have you or anyone you know well had severe illness?[b]	
Transient	18 (53)
Chronic	20 (59)
Deadly outcome	23 (68)
No response	0
How do you view your own health?	
Very good/Good	19 (56)
Just fine	13 (38)
Bad/Very bad	1 (3)
No response	1 (3)

[a] Age was given in one of the listed age-brackets.
[b] Categories are not mutually exclusive.

Various factor solutions were computed with varimax rotation, and a preferred factor solution of four factors was selected based on the interpretability of the factors as well as their statistical features. The factor analysis resulted in four general viewpoints on severity. Across the different viewpoints, there were few areas of consensus on what severity means. The factor descriptions are based on the composite sort for each factor, rather than any one Q sort provided by the participants. These composite sorts represent a weighted compromise of all the exemplars for that factor, i.e., participants with a statistically significant factor loading on that factor and >50% of their communality corresponding to that factor.

Factor I: 'Natural lifespan'

This factor is centred on the notion of a natural course of life and that life is finite, and that a good and dignified life is better than a long one. Death is natural, and not necessarily severe. Individuals on this factor tend to believe that illness is more severe when it affects young people, though severity is not always a function of age. A core aspect of what makes illness severe is how it affects one's ability to live life on one's own terms. A good end to life makes death less severe, while being kept alive involuntarily is severe. Those associated with this factor also tend to reject pain as severe: pain is considered part of life and does not necessarily define severity. This is a distinguishing feature of this factor. Seven out of nine participants who loaded significantly on Factor I were flagged.

Factor II: 'Severity is subjective'

In this factor, severity is determined by the individual and what they consider severe. Severity cannot be defined at the group level, but rather depends on the subjective experience each patient has of their illness. People on this factor are unwilling to define severity according to almost any specific criteria, and reject almost all potential measures of severity, such as ability to work, quality of life, or risk of death. The only criterion considered relevant is age, with illness considered to be more severe if it affects the young. A distinguishing feature of this factor is parenthood. Statements on the relevance of parental responsibilities are not ranked high, but are rejected in all the other factors. This could be related to the overall view that illness affecting the young (whether as patients or as children of patients) is more severe. All five participants who loaded significantly on Factor II were flagged.

Factor III: 'Objective measures and triage'

This factor is centred on objective measures of illness and is reminiscent of a 'medical triage' perspective on severity. Issues such as urgency and prognosis—which are more objectively measurable—are considered to be important in determining severity. As in Factor I and Factor

II, there is an emphasis on age. There is support for the idea that healthcare personnel have a central role in defining severity. Dignity, which is important in all other factors, is ranked distinguishingly low, possibly because dignity is difficult to measure and therefore considered too subjective. The view that all lives are equal, and that severity is not related to status or success, is central in this factor. Four out of eight participants who loaded significantly on Factor III were flagged.

Factor IV: 'Functioning and quality of life'

In this factor, functioning and quality of life are central. There is a certain degree of subjectivity in determining severity, but subjective experiences of severity are not beyond measurement. Severity is defined by the quality of life and by how it affects day-to-day life: whether you can live life as you want, your functioning, and whether you can still enjoy your hobbies, the latter of which is distinguishing for this factor. Dignity and being able to take care of oneself is also important. A distinguishing feature is the rejection of age as a measure of severity. The notion that next-of-kin or parental responsibilities affect the degree of severity is also rejected more strongly compared with other factors. Nine of twelve participants who loaded significantly on Factor IV were flagged.

5.3 The distribution of viewpoints on severity

The aim of Study III was to explore the prevalence and distribution of the factors on severity uncovered in Study II across the Norwegian population. The corresponding paper is titled, 'Charting public views on the meaning of illness severity'. In the following discussion of the results from Study III, the four vignettes (based on the viewpoints, or factors, from Study II) are referred to as 'Lifespan', 'Subjective', 'Objective', and 'FQoL'.

Sample and exclusion criteria

The pilot with SEVPRI's Advisory Board and the pilot with members of the Akershus University Hospital user panel and colleagues led to minor edits in the wording of the vignettes and of the survey. The two Norstat pilots were conducted between December 2022 and January 2023, with 99 and 134 respondents, respectively. Analysis of these demonstrated a high level of inconsistency between respondents' ranking and scoring of the vignettes, meaning respondents would often change the order of the vignettes from the order they first provided in the ranking subtask, to the order they gave the vignettes when scoring them on the VAS. This was thought to be caused by a suboptimal digital solution which made it difficult to place vignettes onto the VAS. Following the two rounds of pilot data collection, and improvement of the digital solution for the vignette scoring task, inconsistencies decreased from approximately four to two out of ten. The final version of the survey was administered from March to April, 2023.

In total, 1174 respondents completed the survey. Comparing the sample to data from Statistics Norway regarding representativity, there was an overrepresentation of women and highly educated individuals, as typically observed in survey studies. Beyond this, the sample was considered adequately representative. Preliminary analysis demonstrated higher levels of inconsistencies (i.e., changing the order of the vignettes between the ranking and the scoring subtasks) for respondents who completed the survey in under 3.5 minutes. Myself and MB also tested the survey and could not complete it at this speed, despite our familiarity with the survey. Wishing to exclude respondents who were unlikely to have meaningfully engaged with the survey, the exclusion criteria of 'completion ≤3.5 minutes' was set, resulting in a final sample of 1094 respondents. Of these, 84.9% provided consistent responses to the tasks. A further 6.9% did not rearrange the vignettes from the randomised order they were presented in. Preliminary analyses explored how respondents had used the VAS in the scoring task. The average range used (from highest to lowest score) was 6. The average highest score was 8.9 and the average lowest score was 3.0, and 62.2% of all scores were 6 or above. Of the sample, 84.7% provided a highest score to one vignette only, singling this out as most like their view.

Vignette alignment

Using the first approach to vignette alignment, placing all respondents onto their highest scored vignette and allowing ties, 40.2% of respondents scored Lifespan highest, 32.4% for FQoL, 28.9% for Objective, and 16.3% for Subjective. Lifespan and FQoL were the most frequent tied pair, while Subjective and Objective were the least frequent tied pair. A two-way tie was observed for 13.0% of participants, with 2.3% providing a tie for three or four vignettes. In the second analysis, adding a requirement of a score ≥7 but still allowing ties, 28.5% of respondents scored Lifespan highest, 22.1% for FQoL, 22.0% for Objective, and 10.9% for Subjective. 1.9% did not score any of the vignettes ≥7. In the third analysis, without the requirement of scoring above a certain point on the scale but resolving ties by referring to the ranking subtask, 34.7% of respondents scored Lifespan highest, 27.0% for FQoL, 25.0% for Objective, and 13.3% for Subjective. The picture remained similar for the strictest approach, representing vignette membership, with the requirement of a score ≥7, excluding ties, and with a gap requirement. We observed a slight shift in support for Objective, moving this ahead of FQoL. In the fourth analysis, approximately 50% of the sample were not assigned vignette membership.

Using different approaches to describe vignette alignment, there appeared to be support for all four vignettes, with all vignettes maintaining some level of support throughout the analyses. Most respondents scored more than one vignette in the 7–10 range, and only 21 respondents did not score at least one vignette in this range. Lifespan was ranked highest across all analyses, and Subjective was ranked lowest. Linear regression analysis suggested that the variables of sex, gross household income, and education level impact vignette alignment, with men, individuals with low education, and individuals with low gross household income more likely to score Subjective high than any other individuals or groups of individuals.

6. Discussion

In the following discussion, I will adopt a broader perspective on the results of the three studies presented in Chapter 5. First, I summarise the findings across the three studies and consider similarities and differences between the findings. Following this, I compare views identified on the meaning of severity with current policy operationalisations. I then discuss the findings in the context of existing empirical work on public views on severity and the literature on severity and priority setting more broadly, before considering if the findings fit into a framework for understanding and categorising public views on severity. Next, I consider severity as a polyvalent concept. Finally, I consider the options facing policymakers when there is a dissonance between public views and policy, before reflecting on potential policy implications of the thesis.

6.1 Public views on the meaning of severity: findings from three studies

In this thesis I explore severity, seeking to ascertain and characterise public views on severity in a healthcare context, and make several important findings. While the three studies in this thesis have different objectives, they all provide knowledge on public views on severity. The thematic analysis of the group interviews in Study I demonstrates that there are several conceptualisations of severity with considerable differences between them, and goes on to suggest that severity is a polyvalent concept. The Q methodological factor analysis in Study II led to the identification of four different viewpoints on severity. Results from the survey distributed for Study III demonstrate that these viewpoints, translated into vignettes, are distributed widely across the population. There appears to be support for all four vignettes, with none singularly representing what severity generally is taken to mean, and with the majority of respondents appearing to align themselves with at least one vignette. This supports the findings in Study I that severity is not ‘one thing’, but rather perceived of as representing a variety of different issues.

Before continuing onto further discussion of the results, I will provide a brief consideration of the various findings across the three studies. First, while the studies identify different views and conceptualisations of severity, they all provide different answers to a similar question: what does severity mean to members of the public? Study I approaches this question as openly as possible, exploring what participants in the group interviews were really talking about and saying when asked what severity meant to them, resulting in different conceptualisations of severity. Study II approaches the question in a more deliberate manner, presenting participants with specific statements about severity and asking them to rank them, ultimately searching for general viewpoints by identifying areas of consensus and disparity on severity. Study III provides a quantitative exploration of how those viewpoints, via vignettes, are distributed across the wider population. Together, the studies provide a broad, explorative, and comprehensive attempt at answering the question of what illness severity means to the public.

The notion of severity as subjective and determined by the individual emerged both in Study I, based on data from group interviews, and in Study II, based on data from the Q sort exercises. Central to both these studies was the notion of severity as subjective self-evaluation, centring severity on individual experience and the individual's own perception of their situation. However, in the subjective conceptualisation of severity in Study I, any form of outside, objective evaluation of severity appeared to be rejected, while in the subjective viewpoint in Study II, the relevance of age as a criterion for determining severity emerged as relevant for some participants.

Severity as an objective notion could also be identified in both Study I and Study II, with the idea that severity is determined by certain objective criteria pertaining to a condition. In the objective conceptualisation in Study I there was no general agreement on what these criteria might be, with emphasis rather on the idea of severity as objective fact. In Study II, specific criteria emerged as relevant objective measures of severity, including urgency, prognosis, and age. As a note, a striking finding from Study III was the particular antagonism between the vignettes representing the subjective and objective viewpoints on severity: these were the least commonly tied vignettes, had the fewest common supporters, and few respondents scored them both either high or low. This is theoretically reassuring, as they represent directly opposing views on severity, and suggests respondents understood the vignettes the way we intended them to.

In the situational conceptualisation of severity in Study I, three subthemes were identified, all centred on the effects of illness. In the subtheme of social effects of illness, emphasis was on how illness affects one's functioning and capabilities, and ability to lead a normal life. This bares comparison to findings in Study II, considering the 'natural lifespan' factor, with emphasis on reaching important milestones and leading the life one wishes to, and the ‘functioning and quality of life’ factor. Considering the relevance of the effects of illness on those surrounding the patient and effects at a societal level, evident as a theme in the group interviews and central to the situational conceptualisation of severity in Study I, did not appear in the factor analysis for Study II, and is therefore not represented elsewhere in the data. The notion of fairness as relevant to determining severity—either for the individual, within and across patient groups, or for society—was also unique to Study I.

Findings from Study III suggest that, while the majority of respondents across the four categories of analysis scored ‘Lifespan’ highest and ‘Subjective’ lowest, there was some level of support for all vignettes, and no single vignette can be said to adequately represent what the citizenry take severity to mean. The findings further demonstrate that respondents generally tended to agree with at least one vignette, and often more than one, and used the top half of the VAS more often than the bottom half. This suggests that respondents felt several of the vignettes represent their views on severity. This supports claims made in Study I and Study II that severity is a complex concept with many different meanings associated with it—even by the same individual.

Overall, the conceptualisations and viewpoints identified in Study I and Study II and the responses to the vignettes in the Study III can be considered to represent the breadth of public views on severity. Severity appears to be characterised by plurality, with a variety of different views on the meaning of severity—many of which appear to have conflicting emphases.

6.2 Comparing findings to policy operationalisations of severity

Considering the various views we identify on the meaning of severity among members of the Norwegian public, a central question is how these views compare to policy operationalisations of severity. The following comparison is based on the conceptualisations and viewpoints identified in Study I and Study II. Note that the vignettes in Study III are based on the factors from Study II and summarise the full viewpoints, and this discussion considers the viewpoints in their rich and detailed form.

While the findings in this thesis originate from a Norwegian context, I compare the views we identify both to the absolute QALY shortfall operationalisation of severity in the Norwegian priority-setting framework and to operationalisations in other jurisdictions. The reasoning behind this is that there is an ongoing discussion in the Norwegian healthcare context on how severity ought to be operationalised, and while absolute QALY shortfall represents the current operationalisation, other versions have also been put forth. Considering how views of the Norwegian population compare to other severity operationalisations is therefore of value for policymakers, especially as these operationalisations are applied in publicly-funded healthcare systems which are comparable to the Norwegian model.

As stated in Chapter 2.3.1, severity is operationalised in Norway as absolute QALY shortfall (Ministry of Health and Care Services, 2020); in the Netherlands as proportional QALY shortfall (Reckers-Droog et al., 2018); and in the UK as a mix of the two (National Institute for Health and Care Excellence, 2022). The Swedish operationalisation is based on a descriptive version of different severity levels (Riksdagsförvaltningen, 2018). I begin by comparing our findings to the Norwegian operationalisation, where severity is based on the principles of risk of death or loss of functioning, degree of physical or mental loss of functioning, and pain, physical or mental discomfort, and quantified by absolute QALY shortfall estimations.

Several of the identified views appear to be partly accommodated by the Norwegian operationalisation. Both the objective conceptualisation of severity in Study I and the ‘objective measures and triage’ factor in Study II emphasised that severity should be determined equally and objectively within and across patient groups. This is closely aligned with an important intention behind QALY-based approaches to severity (in both absolute QALY shortfall and proportional QALY shortfall estimations), i.e., to ensure objective and fair resource distribution (Spencer et al., 2022). The emphasis on quality of life, as seen in the ‘social effects’ conceptualisation in Study I and in both the ‘natural lifespan’ and ‘functioning and quality of life’ factors in Study II, is also closely tied to QALY estimations, which quantify both quantity and quality of life.

Some views appear to align specifically with an absolute QALY shortfall operationalisation. While age is not emphasised outright—though, it has been argued, indirectly (Tsuchiya, 2000)—in absolute QALY shortfall estimations, illnesses affecting younger patient groups have a greater absolute QALY loss in absolute QALY shortfall estimations compared to their relative QALY loss in proportional QALY shortfall estimations. Age was considered relevant (though emphasised to different degrees) in the ‘natural lifespan’, ‘severity is subjective’, and ‘objective measures and triage’ factors in Study II. These factors thus compare to outcomes of an absolute QALY shortfall-based severity operationalisation. Similarly, the particular emphasis on mental illness as expressed in the 'natural lifespan' factor also bares comparison to the current Norwegian operationalisation. However, many of the views we identified contrast with absolute QALY shortfall outcomes. For example, the rejection of both pain and death as particularly severe in the 'natural lifespan' factor directly contrasts the Norwegian operationalisation of severity. Stigma, considered relevant in the 'severity is subjective' factor, is not accounted for in the current operationalisation. And while some of the views we identified are better accounted for in a proportional QALY shortfall operationalisation, including the rejection of age as relevant to severity in the 'functioning and quality of life' factor in Study II, many of our findings do not seem reconcilable with a QALY-based approach. In the subjective conceptualisation of severity in Study I and the comparable 'severity is subjective' factor in Study II, emphasis was on the inherently subjective nature of severity and the right of the individual to determine severity for themselves, rejecting the use of criteria. This is in stark contrast to the principles behind QALY-based approaches to severity, and it is difficult to see how such views could be accommodated in a QALY-based framework. Further, the emphasis within the situational conceptualisation of severity on how those surrounding the patient and society are affected by illness are explicitly excluded in national guidelines with both absolute QALY shortfall and proportional QALY shortfall operationalisations (National Institute for Health and Care Excellence, 2022; NOU 2014: 12, 2014).

While the Swedish severity framework has a text-based operationalisation, severity is estimated via measures similar to QALY shortfall outcomes. This includes considerations of loss of quality of life, the duration of illness, future ill health and risk of death, and remaining life years (Broqvist, 2018; Riksdagsförvaltningen, 2018). As such, few of the additional concerns we identified appear to be accommodated by the Swedish operationalisation. There is, however, particular emphasis on loss of functioning, comparable to concerns in the 'natural lifespan' and 'functioning and quality of life' factors in Study II and the situational conceptualisation of severity in Study I. Unique to the Swedish operationalisation, there is also emphasis on the role of participation, with illness considered more severe if it negatively impacts one's ability to participate in different aspects of everyday life. We identified this concern in the 'social effects of illness' subtheme in Study I. Notably, however, both QALY-based operationalisations and the Swedish model are patient-centric, with severity determined by the effect illness has on the patient. The relevance of the effects on those surrounding the patient and on society, as emphasised in the situational conceptualisation of severity in Study I, appears to be absent in current operationalisations of severity.

As outlined in Chapter 2.3.1, the Norwegian operationalisation of severity contains both a group-level, quantitative version of severity (absolute QALY shortfall) as well as a descriptive version intended for the clinical level. Notably, the latter includes several of the additional concerns we identify in our material. To illustrate, I include an excerpt from the Magnussen working group's report on what the descriptive approach to severity is intended to represent:

> "Examples of aspects that make a disease and the overall situation more severe could be whether the condition without healthcare services puts the patient in an undignified situation, whether the patient regardless of the condition is particularly disadvantaged or whether the condition leads to reduced work ability or reduced social functioning for the patient and the communities the patient is a part of […] Our point is that such circumstances are in fact at times taken into account in a clinical setting – individual patient are given higher priority because there are certain conditions that make their illness and situation more severe. This will sometimes be uncontroversial, while other times it will be controversial and set basic conceptions of fairness up against each other. For example, is it right to prioritise the working population over the unemployed, disabled and pensioners? The working group does not take a position on such questions, but in principle will give its support to clinicians having the opportunity to assess whether particular aspects of the patient group, patient and situation make the patient's illness and overall situation particularly severe—and give the patient priority accordingly." (Magnussen et al., 2015, p.21-22) [Translated from the original Norwegian language version]

This descriptive version of severity accounts for many of the additional concerns our participants considered relevant to severity, which are not accounted for in the absolute QALY shortfall iteration of the Norwegian severity criterion, nor in proportional QALY shortfall estimations or in the Swedish model. First, the relevance of the individual situation and taking the individual's context into consideration compares to the subjective conceptualisation and the 'severity is subjective' factor. The emphasis on the unique insight of clinicians is recognisable from the objective conceptualisation and the 'objective measures and triage' factor. The text also points to the relevance of dignity, as emphasised in both the 'natural lifespan' and 'functioning and quality of life' factors. Finally, this descriptive version is the only operationalisation which mentions the potential relevance of the effects of illness on next-of-kin, as well as the societal effects of illness, such as costs and productivity. Such a societal perspective is central to the situational conceptualisation of severity identified in Study I, where severity was conceptualised by what the costs and effects of illness are on society, rather than only on the patient or the healthcare system. While the working group considers all these concerns "morally relevant" (Magnussen et al., 2015), they state that they will not take a position on these questions. Rather, they provide clinicians with the opportunity to assess when and to what degree these concerns ought to confer priority. However, it is unknown to what degree these concerns are included in priority-setting processes at the clinical level. We therefore cannot know that the views we identified are in fact represented in any systematic way in priority-setting decisions.

In the group interviews, an issue that participants persistently tied to severity was the matter of fairness. As the findings of Study I suggest, participants appeared to consider it important that severity is determined in a fair manner, though there were differing views on what constitutes fair. Some of the approaches to fairness, as participants related it to severity, are comparable to central principles in healthcare policy, such as the value of equal access and reducing health inequalities (Commission on Social Determinants of Health, 2008; Williams, 1988). However, results from Study I illustrate other conceptions of fairness as well, which contrast—some starkly so—notions of fairness established in healthcare (Olsen, 2011a). In the subjective conceptualisation of severity, there was an emphasis on determining severity fairly according to the individual's subjective viewpoint. This view of fairness in relation to severity, centred on the value of lived experience and the individual's unique understanding of their illness, compares to a common critique of basing QALY estimations on public preferences rather than patient perspectives (Whitehead & Ali, 2010). It also contrasts the concern for fairness in the situational conceptualisation, which aligns with one of the arguments in support of public preference-based QALY estimations: that, as the public both pays for and is the recipient of healthcare, they ought to have a say in how healthcare is distributed (Whitehead & Ali, 2010).

Within the situational conceptualisation we also identified contrasting notions of fairness to those commonly established in priority-setting frameworks. These frameworks typically centre on ensuring fair distribution within the healthcare sector, for the patients, rather than beyond the healthcare sector, for society (Daniels, 2007). Finally, within the objective conceptualisation, concern for fairness entailed that severity should be determined in an objective manner to avoid punishing those who adapt well to illness. This view compares to the notion of actuarial fairness and the idea that one should not be punished for one's efforts (Olsen, 2011a), resembling a libertarian concern for the better-off (Grossman, 2013). This represents an antithetical approach to fairness to the egalitarian concern for the worse off entrenched in publicly-funded healthcare systems (Brock, 2002a; Ministry of Health and Care Services, 2016; Nord, 2005; Norheim et al., 2019; The Standing Committee on Health and Care Services, 2022).

It is evident that severity is an ambiguous concept both in the literature and in policy. Our findings demonstrate that severity appears to be equally ambiguous in the public sphere, and several of the public views we identify are unaccounted for in current policy operationalisations of severity. It appears that QALY shortfall operationalisations fail to capture many of the common intuitions on severity held by members of the public, pointing to a misalignment between public views and policy operationalisations of severity. This contradicts the claim in Norwegian policy documents that an absolute QALY shortfall operationalisation expresses "society's approach to severity" (Ministry of Health and Care Services, 2016). With the plurality of views identified, the findings of this thesis suggest that there is no one view of severity held by members of the public, and that "society's approach to severity" is in fact not one thing.

6.3 Comparing findings to existing literature

In the following section, I consider our findings in light of previous empirical work on public views on severity. As outlined in Chapter 2.4.2, there are multiple studies exploring whether or not severity is considered relevant to priority setting in healthcare (Gu et al., 2015; Shah, 2009; Skedgel et al., 2022). They establish that citizenries across multiple countries consider severity to be an important, if not the most important, consideration in priority setting. Overall, the studies addressing public views on severity analysed in this thesis demonstrate that public views on severity is considered an interesting and important research question. However, across these studies, severity is often left undefined, and if it is explained (either for the reader or participants) severity is defined in a variety of ways, including definitions based on quality of life, life expectancy if untreated, pre-treatment health problems, and proximity to death. Often, the authors across these studies presuppose a certain definition of severity, e.g., health-related quality of life, and when their respondents provide a viewpoint regarding the relevance of health-related quality of life to priority-setting, the authors then assume they have provided a viewpoint on the relevance of severity. We cannot know if that is what the respondents have in fact provided. Considering the variation in both design and approach to severity, comparing the results from these studies to each other and to the results presented in this thesis, is difficult. Further, while these studies demonstrate that members of the public consider severity to be an important concern in priority setting, they do not provide answers to, or in fact question, what members of the public take severity to mean. These studies appear to presuppose different definitions of severity, and while they demonstrate that members of the public are generally in favour of prioritising treatment for the severely ill, we do not know which patient groups respondent in these studies consider to be severely ill.

To my knowledge, the only comparable studies exploring public views on the meaning(s) of severity in the context of healthcare are those conducted by Broqvist and colleagues and by the Magnussen working group. Our findings demonstrate that the public have multiple, contrasting approaches to severity, and that severity appears to be a complex concept with many different attributes. This supports the findings of Broqvist and colleagues, who describe severity as a complex and multifactorial concept (Broqvist et al., 2018). They compare their findings to the Swedish severity framework and find that many of the attributes the citizenry associate with severity are covered by the framework, while some are not. Comparing their findings to a QALY shortfall operationalisation, it appears that some attributes are covered by the current Norwegian operationalisation of severity, such as functioning, risk of death and future ill health, and duration of illness. However, participants in their study also point to potential attributes of severity which are not included in QALY shortfall operationalisations, such as restrictions on participation and social life imposed by illness, societal costs, and the patient's own ability to affect the impact of illness on their lives. These attributes, although excluded from policy, compare to the findings in this thesis. The role of participation and being able to live one's life is reflected in the situational conceptualisation of severity in Study I, as well as in the 'natural lifespan' and 'functioning and quality of life' factors in Study II. The issue of societal costs was a central concern in the situational conceptualisation in Study I, while the issue of ability and effort to rehabilitate from illness was addressed in light of fairness in the objective conceptualisation in Study I. Overall, it appears that all the potential attributes of severity in Broqvist and colleagues' study appear in the findings of this thesis, while the thesis also identifies additional concerns, such as subjective and objective notions of severity and the role of fairness in understanding severity.

The findings in this thesis also support results from the questionnaire distributed by the Magnussen working group among policymakers, healthcare workers, and patient organisations. These participants were sought out specifically for their healthcare background, but are, nonetheless, also members of the public. Based on their questionnaire, the Magnussen working group suggest that there is no uniform agreement on what severity means, with participants not only describing severity with different words and formulations, but also with different content and meaning (Magnussen et al., 2015). Considering Broqvist and colleagues' and Magnussen and colleagues' studies, which appear to be the only other empirical works exploring public views on the *meaning* of severity, it seems that they all corroborate the claim that severity is complex, multifaceted, and more than QALY shortfall operationalisations account for. The findings in this thesis suggest that when asking members of the public about their views on severity, they do not all take severity to mean the same thing. This is an important contribution that can be built on in further work that seeks to elicit public views on severity.

A subsequent consideration is whether these views can be located within the literature on severity specifically or on priority setting and healthcare generally. There are many parallels to be drawn. The objective conceptualisation in Study I and the 'objective measure and triage' factor in Study II compare to familiar positions in the literature. The perhaps most notable parallel is to criteria-driven approaches to healthcare centred on health evaluation via objective, evidence-based standards (Cleary, 1997; Marsh et al., 2014; Mobbs, 2021). The contrasting approach to severity represented by the subjective conceptualisation in Study I and the 'severity is subjective' factor in Study II are also recognisable in the wider literature, comparable to phenomenological approaches to illness and the importance of lived experience (Reynolds, 2022; Toombs, 1995). These contrasting perspectives—between the impersonal and objective versus the personal and subjective—are also familiar positions in the literature on severity in genomics, where there is an ongoing debate on which perspective best describes the nature of severe (interchangeably using the terms 'severe' and 'serious') genetic illness (Boardman & Clark, 2022; Dive et al., 2023; Newson & Dive, 2021).

The 'severity and the social effects of illness' subtheme from Study I, considering severity according to one's ability to work, maintain relationships, and participate in activities, and the ‘natural lifespan’ factor in Study II, connecting severity to reaching important milestones and living the life one wishes to, bear comparison to the capability approach. This framework is central in public health ethics and emphasises the value of capabilities and functioning (Nussbaum & Sen, 1993; Prah Ruger, 2010; Sen, 1993). Capabilities are understood as the opportunity to lead the life one wishes to, including the opportunity to achieve one’s goals and reach important milestones. Functioning is understood as the realisation of those capabilities, such as employment, maintaining relationship, enjoying hobbies, etc. (Robeyns & Byskov, 2023; Venkatapuram, 2013). ‘Natural lifespan’, with emphasis on reaching important milestones, also resembles the theory of health put forth by Swedish philosopher Lennart Nordenfelt, based on health as the ability to achieve vital goals (Nordenfelt, 1995; Venkatapuram, 2013). This factor is also reminiscent of the ‘fair innings’ argument in health economics, based on the notion that everyone is entitled to a certain amount of health and that everyone deserves to reach a certain threshold of health and (quality-adjusted) lifespan (Williams, 1997b). The ‘functioning and quality of life’ factor in Study II also compares to the capabilities approach, as well as the literature on the relevance of quality of life to illness severity (Hofstede, 1984; MacKillop & Sheard, 2018; Nord & Johansen, 2014; Ottersen et al., 2016; Salomon, 2017).

The ‘severity and the effects of illness on others’ and ‘severity and the effects of illness on a societal level’ subthemes in Study I represent a departure from the healthcare perspective that the majority of the identified views hold, which centre on severity understood by the effects of illness on patients. In ‘severity and the effects of illness on others’ and ‘severity and the effects of illness on a societal level’, however, severity was conceptualised through effects on entities beyond the patient. The notion that severity depends on the effects on next-of-kin compares to a concern for spillover effects (Al-Janabi et al., 2016; Canaway et al., 2019; Gustavsson et al., 2023). Concern for the effects of illness at a societal level can be interpreted to align with the notion that resource distribution ought to be determined partly by social productivity, prioritising the ‘productive citizenry’ (Dineen, 2011; Schneiderman, 2011). This is in stark contrast to the well-established principle in health economics literature that contribution ought to be divorced from need (Olsen, 2011b). The emphasis on productivity further compares to the societal perspective in health economic literature more widely (Drost et al., 2017; Drummond et al., 2015).

These approaches to severity, anchored in concerns beyond the patient, speak to a wider, ongoing debate in healthcare priority setting on whether priority-setting frameworks ought to have a healthcare and/or societal perspective, and whether costs outside the healthcare system, such as production loss, ought to be included in health economic evaluations. The Norwegian government has recently appointed three expert groups to develop a 2024 white paper on priority setting (Ministry of Health and Care Services, 2023b). Demonstrating the currency of this dilemma of healthcare versus societal concerns, one of these groups—the 'Perspective group'—is commissioned to address whether the current healthcare perspective should remain the foundation for priority-setting decision in the Norwegian healthcare system, or whether a broader societal perspective ought also to be included.

6.4 A framework to accommodate the polyvalence of severity

Considering the many issues that members of the public associate with severity, a detailed analysis could be undertaken to continue to explore how these compare to positions in the literature. However, in light of the plurality of views we have identified, considering whether these views can be placed within some form of framework, providing a more holistic understanding of public views on severity as we have identified them, is perhaps a more helpful exercise for policymakers, for whom public views on severity are of particular interest. Across the many views on severity, are there any commonalities that allow for some form of categorisation of these findings?

In Study II, the four factors are briefly explored in light of Twaddle's Triad, analysed against the three perspectives of illness, sickness, and disease (Hofmann, 2002; Stenmarck et al., 2023; Twaddle, 1994). The three concepts within this triad reflect different perspectives on human ailment. However, there appear to be some parallels between descriptions of the perspectives in Twaddle's Triad and the various views we uncover on severity. I therefore explore whether this framework lends itself to a categorisation of public views on severity, as they have been identified across the three studies in this thesis.

The first of Twaddle's perspectives is *disease*. This represents a biomedical view on ailment, typically as conceived of by healthcare professionals (Hofmann, 2002; Twaddle, 1994). Disease is considered independent of the individual's subjective experience and can be measured objectively (Hofmann, 2002). This perspective is clearly recognisable within the objective conceptualisation of severity in Study I, where severity was conceptualised via the use of objective criteria. It is also comparable to the 'objective measures and triage' factor in Study II, viewing severity through a medical triage approach.

The second perspective in the triad is *illness*. This perspective is tied to subjective experience and how illness is conceived of by the individual (Hofmann, 2016; Yew & Noor, 2014). According to the triad, illness "is a subjectively interpreted undesirable state of health. It consists of subjective feeling states (e.g., pain, weakness), perceptions of the adequacy of their bodily functioning, and/or feelings of competence" (Twaddle, 1994). This perspective aligns with the subjective conceptualisation in Study I, centred on the individual's experience, and the 'severity is subjective' factor in Study II, oriented around severity determined by the individuals own view of their condition. In both, severity was considered to depend on the individual's evaluation of their situation.

The third and final perspective in the triad is *sickness*, which centres on the social phenomena of ailment (Friedman, 2021; Hofmann & Wilkinson, 2016). Within the triad, sickness is also tied to social identity, and—on one account—how sickness affects the ability to function in society. Several of the identified public views on severity align with this perspective. The emphasis on the social effects of illness in the situational conceptualisation of severity in Study I, and how severity is related to one's ability to partake in relationships, hobbies, and work, aligns with the sickness perspective. The 'natural lifespan' factor in Study II also compares to this perspective, emphasising the value of living a full life and reaching important milestones, as does the 'functioning and quality of life' factor. A central aspect of the sickness perspective, however, centres on how ailment is perceived by and affects society. This compares to the 'severity and the effects of illness on others' and 'severity and the effects of illness on a societal level' subthemes in Study I, where severity was conceptualised by how a condition affects those surrounding the patient and effects at a societal level. An important difference between the sickness perspective within the triad and the situational conceptualisation of severity, however, is that the triad emphasises how society's perception of sickness entitles *patients* to treatment, economic support, and freedom from obligations like work (Hofmann, 2016). In the situational conceptualisation in Study I, emphasis was on determining severity by considering costs of illness on *society*, and the higher the cost for society, the more severe the condition.

It seems that the commonalities and distinctiveness between the views we identified allow for a consideration of severity along the perspectives presented in Twaddle's Triad. However, severity appears to mean different things to different people, and, as pointed out in Study I, severity also appears to mean many things at once. While Twaddle's Triad distinguishes between disease, illness, and sickness as three separate perspectives on human ailment—albeit with some overlap (Hofmann, 2016)—it appears that, in the case of severity, these categories are perhaps best described as different qualities of severity, rather than different perspectives on it (Friedman et al., 2022). While the perspectives presented in Twaddle's Triad do not provide a perfect fit for views held by the public on severity, it seems they present a helpful roadmap in the very complex territory that the plurality of public views on severity represent.

The studies in this thesis have different aims: Study I is a qualitative, in-depth exploration of how severity is conceptualised by members of the public; Study II represents a mixed-methods approach to identifying the breadth of views on severity and exploring areas of consensus and dissent between them; and Study III reports on a quantitative survey, where the distribution of vignettes (based on the viewpoints from Study II) across the Norwegian population is described. However, despite the differences in the aims of these studies and the methodologies applied to meet them, they all point to a central, important finding:

Severity, for members of the public, is not* one *thing.

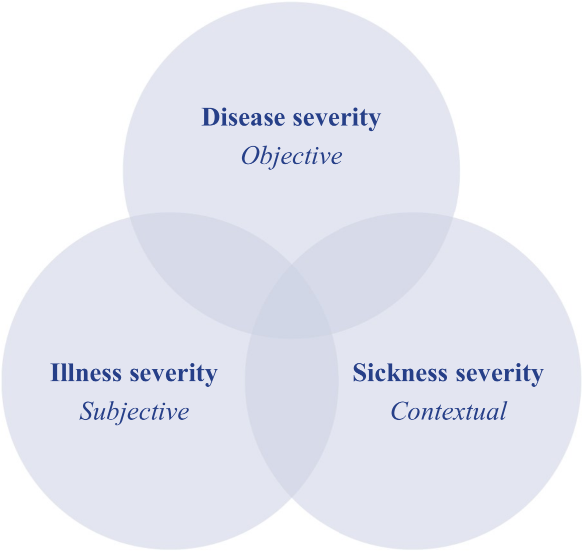
The plurality of views on severity suggests that the observation made in Study I rings true across all the findings on severity this thesis has produced: severity is a polyvalent concept (Pilgrim, 2008). There are many differing views and assumptions underpinning the meaning of severity in a healthcare context, and no single view can describe what severity is taken to mean among members of the Norwegian public. This complexity was perhaps best summed up by one of the conversation participants:

> "I feel severity, it's so ambiguous. Depending on who is considering it, what situation you're looking at it from, what case you throw severity into…if it's death, if it's illness, if it's treatment… It's a word that, yes, seems like some kind of common denomination for everything that we just get lost in."

It is perhaps unsurprising that participants across our studies struggled to determine what severity means, as severity appears to be conceptually complex. Participants approached severity both as a descriptive term, using it to provide observations and characterisations of different conditions, and as a subjective, value-laden concept. This suggests severity is both a descriptive and a normative concept (Schoon & Chi, 2022). Both these aspects of severity are evident in the material of this thesis. The objective conceptualisation of severity in Study I and the 'objective measures and triage' factor in Study II correspond to a descriptive understanding of severity, looking at specific qualities of severity that determine whether a condition is or is not severe. The subjective conceptualisation of severity in Study I and the 'severity is subjective' factor in Study II, on the other hand, appear to represent normative approaches to severity, considering severity from the point of view of the individual and what they experience as severe or not. The situational conceptualisation of severity in Study I and the 'natural lifespan' and 'functioning and quality of life' factors in Study II contain both descriptive and normative qualities. This ties severity to the notion of *thick concepts* (van der Weele, 2021), which involve both evaluative and non-evaluative descriptions (Väyrynen, 2021). As with other thick concepts, severity appears to have multiple dimensions of content and meaning associated with it (Thacher, 2015).

It appears that the polyvalence of severity can partly be explained by it being a thick concept with both descriptive and normative features. Returning to the characterisation of the findings along the perspectives provided by Twaddle's Triad, which in the case of severity are perhaps more accurately described as different qualities of it, I venture that a second explanation for the polyvalence of severity can be tied to these different qualities. While severity is discussed in the literature and operationalised in policy as one concept, for the citizenry, severity appears to be threefold. Our findings suggest that severity, according to the views of the public, consists of three elements, which can be described as illness severity, disease severity, and sickness severity (see Figure 3).

Figure 3 The three central qualities of severity.



These elements can be seen as three partly overlapping spheres, representing distinct but interlinked qualities. Disease severity represents the medical, objective, criteria-driven qualities of severity, as observed in the objective conceptualisation in Study I and the ‘objective measures and triage’ factor in Study II. Illness severity represents the subjective, phenomenological qualities of severity, as identified in the subjective conceptualisation and the ‘severity is subjective’ factor. Finally, sickness severity represents the contextual qualities of severity, based on how a condition affects one’s ability to function and enjoy life. These qualities can be identified across the findings, but perhaps most evidently in the ‘natural lifespan’ and ‘functioning and quality of life’ factors in Study II, as well as in the subtheme of the situational conceptualisation in Study I centred on the effects of a condition on work, hobbies, relationships, etc. Sickness severity also represents the broader context surrounding a patient, identified as the relational and societal effects of a condition in the situational conceptualisation. The polyvalence and conceptual complexity of severity could thus in-part be explained by these distinct qualities—all of which appear to be perceived as qualities of severity by the citizenry.

6.5 Addressing misalignment between public views and policy

I have demonstrated that there is a misalignment between public views of severity and operationalisations of severity in policy. A next consideration is what policymakers can do with this knowledge, and how they might address the plurality of views within the public and the subsequent inevitable misalignment between some of these public views and policy. At the outset, it is worth considering why this misalignment matters. First, there is the matter of democratic legitimacy. Involving the public in policy-making and decision-making is democratic, and the public therefore ought to be given the opportunity to participate and be heard (Baker et al., 2021). In Chapter 2.4.1, I present arguments supporting the relevance of public views to policy formulation in ensuring democratically legitimate systems and point to the participatory turn in healthcare generally and priority setting specifically (Abelson et al., 2013; Baker et al., 2021). There is no straightforward answer to how and to what degree public views ought to be implemented in policy-making, and that discussion is beyond the scope of this thesis. However, if policy is in misalignment with public views, and no justification is provided for this misalignment, this arguably jeopardises the democratic legitimacy of that policy. Misalignment could thereby represent a democratic problem for policymakers in and of itself.

Second, the importance of healthcare to the public is well-established in both public discourse studies and in the wider literature (Hannawa et al., 2022; Hirose & Bognar, 2014; National Research Council (US) & Institute of Medicine (US), 2013; Stenmarck et al., 2021; Williamson, 2014). The outcomes of priority-setting decisions have implications for the public whose health, and even survival, might rely on those outcomes. This makes priority setting in healthcare a contentious issue. As healthcare budgets become increasingly strained, the public discourse on priority setting demonstrates tension on the matter of healthcare rationing (Broqvist & Garpenby, 2014; Stenmarck & Nilsen, 2022). Misalignment between policy and public views is likely to increase this tension and contribute to a more contentious priority-setting discourse, with miscommunications between policymakers, decisionmakers implementing those policies, and the citizenry. A dissonance between public views and policy operationalisations on central priority-setting principles will likely make it difficult for the public to understand and support priority-setting outcomes.

Different approaches can be taken to address plurality (Baker et al., 2021; Gutmann & Thompson, 2004) and policymakers have several strategies available to them (Broqvist, 2018). One strategy is to implement public views more directly in policy-making, and policymakers can choose to implement these views in priority-setting decisions via increased responsiveness (Dahl, 2008; Pitkin, 1967; Williams et al., 2014). An example of such responsiveness was demonstrated by Norwegian policymakers when the 2014 Norheim committee suggested that the severity criterion should be reformulated into a health loss criterion (see Chapter 2.3.1). This involved operationalising severity by considering expected lifetime QALY loss (relative to a standard reference life of 80 good life years) rather than future QALY loss, which is used in the absolute and proportional QALY shortfall approaches.

The health loss criterion, however, proved controversial and lead to widespread public debate (Barra et al., 2020; Horn et al., 2021). The subsequent appointment of the Magnussen working group to re-evaluate the severity criterion resulted in the recommendation of an absolute QALY shortfall operationalisation of severity. This real-life example demonstrates the notion of responsiveness in action, and political will and ability to listen to the views (and in this case, protests (Horn et al., 2021)) of the public and to accommodate policy accordingly. Interestingly, when the notion of past health came up in the group interviews, participants appeared to universally reject it as a constituent dimension of severity. This does not mean that participants rejected past health as relevant to priority setting, and it should also be reiterated that the group interviews were conducted with a non-representative sample of the population. However, the data resonate with and reproduce the intuitions of the public surrounding past health that arose in the public discourse following the proposition of the health loss criterion.

A second strategy for handling misalignment between public views and policy relates to transparency, with an educational approach. Policymakers and decisionmakers can work to inform the public about the principles underlying priority-setting decisions and the outcomes these principles translate to in practice (Broqvist, 2018; Williams et al., 2012b). While it is perhaps unfeasible, and unreasonable, to base policy-making entirely on public views and values (Bognar, 2012; Walker & Siegel, 2002), a misalignment between public views and policy arguably requires that the public is provided with the rationales on which policies are made. Esaiasson and Wlezien argue that it is the responsibility of policymakers, and the decisionmakers who implement policies, to provide clear accounts of these rationales (Esaiasson & Wlezien, 2017). Policymakers can also apply the educational approach internally, looking to the experience of other healthcare systems and integrating these in their own. There is ongoing, systematic work with priority setting and establishing priority-setting frameworks in multiple countries, and there are comparable experiences policymakers can learn from to improve the quality and transparency of their own frameworks and systems (Ham, 1997).

A third strategy is to dismiss public views as relevant to policy-making. Such a strategy could perhaps be justified by arguments suggesting that policy ought to be informed by experts rather than lay views, and that priority setting is a complex issue that should be left to those who have expertise on the matter (Broqvist, 2018; Butler & Dynes, 2016). Some participants in the conversations did in fact express views aligned with this strategy, with one participant stating that they believed that "[…] severity is something the professionals within that field should comment on, not the relatives, not the people who are ill". Returning to the two grounds for why misalignment between public views and policy matters—concern for democratic ideals, and tension in the public discourse—it seems that some of these strategies are more appropriate than others. In terms of ensuring democratic legitimacy, ignoring the views of the public and neglecting responsiveness risks leaving political systems in a 'democratic deficit' (Warren, 2009), suggesting some level of responsiveness is appropriate.

As healthcare systems face both increased demand and limited resources, tension surrounding priority-setting outcomes are likely to rise. This places greater demands on priority-setting processes to ensure that the results of these are understandable and acceptable to the public (Schoon & Chi, 2022). Priority-setting outcomes may not always align with public views, especially when these are disparate and characterised by plurality, and cannot all be accommodated. However, ensuring transparency in the operationalisation of priority-setting principles might ameliorate some of the tension in the public discourse on priority setting. If the citizenry is made aware of and can understand the basis for priority-setting decisions, it seems likely that some of the tension surrounding the outcomes of those decisions would be mitigated. This is especially important when policymakers invoke terms used in everyday parlance.

'Severity' is a term commonly used in everyday language and which members of the public can be expected to have varying associations with, as the results presented in this thesis demonstrate. When using everyday terms in policy formulation, and particularly as priority-setting principles, policymakers ought to consider that members of the public are likely to have certain expectations of how those terms translate into policy, and in the case of severity, into priority-setting outcomes.

Ensuring transparency in priority setting is perhaps particularly important at the clinical level, where priority-setting decisions and resource allocation occurs on a daily basis. As outlined in

Chapter 2.3, the term ‘severity’ is commonly used both in everyday clinical discussions and in clinical guidelines (American Psychiatric Association, 2013; World Health Organization, 2022). In a report for the Norwegian Health Directorate exploring priority setting at the clinical level, dialogue with different clinical specialties revealed that clinicians were, overall, unfamiliar with the established priority-setting principles (The Directorate of Health, 2018). When presented with the three priority-setting criteria in the Norwegian healthcare system (presented in Chapter 2.3.1), severity was the criterion that the majority felt best represented what they personally emphasised in their clinical practice, but results from the report also suggest that there was no systematic interpretation of the term or of how severity should be interpreted or applied. This points to an important challenge for priority-setting frameworks: there appears to be a lack of systematic priority setting, with inconsistencies—and thereby lack of transparency—in decision-making across different levels of care, and across different patients and patient groups. While severity is used widely in the clinical setting and in guidelines, there appears to be considerable ambiguity surrounding its interpretation and implementation in clinical practice.

6.6 Policy implications and future perspectives

Healthcare systems globally face a changing reality: sociodemographic developments lead to increasing demands on healthcare systems. Public health challenges, such as pandemics and lifestyle diseases, place new demands on priority-setting frameworks, which increasingly must accommodate both treatment and preventive measures. The move towards precision medicine means treatments are increasingly intended for smaller and smaller groups, challenging the group-level approach health economic evaluations traditionally operate with (Hirose & Bognar, 2014; Ministry of Health and Care Services, 2023b). This changing reality places heavier demands on the robustness and transparency of priority-setting frameworks.

It is my hope that the findings presented in this thesis can be of use to policymakers in future iterations of priority-setting frameworks applying a severity criterion. I have sought to explore public views on severity—but many questions remain unanswered, and new ones have emerged. SEVPRI asked members of the public what they take severity to mean. We have not had the opportunity to explore to what degree the citizenry think severity ought to confer priority in the healthcare system, and if so, which version(s) of severity has support as a priority-setting principle. Further investigation of public views on how severity relates to resource allocation would provide additional and important knowledge for policymakers.

There are many decisionmakers within the healthcare system, and both policymakers at the group level and healthcare workers in the clinic are instrumental to resource distribution. As different operationalisations of severity are provided for the group level and the clinical level in the Norwegian context, future research should examine if views on the meanings of severity are as polyvalent among different decisionmakers and stakeholders as they are for members of the public, and how severity is interpreted and operationalised at the different levels of decision-making.

The findings in this thesis may be of particular value for the abovementioned white paper commissioned by the Norwegian government to address the challenges ahead (Ministry of Health and Care Services, 2023b). The findings of this thesis confirm, as previous research has suggested (Barra et al., 2020; Broqvist et al., 2018; Hausman, 2019b; Magnussen et al., 2015; Olsen, 2013; Solberg et al., 2023), that severity appears to be an ambiguous concept, with lack of any general consensus on it meaning. The studies in this thesis demonstrate that severity appears to be a polyvalent concept, with many different meanings attached to it, and illustrate that there is a plurality of views on severity which are not necessarily reconcilable. In the foregoing discussion, I explored our findings in light of policy, demonstrating that there is a misalignment between public views and policy operationalisations of severity. Considering the ambiguity, polyvalence, and plurality that appears inherent to severity, and the misalignment between public views and policy, two central questions emerge: (i) exactly what is severity intended to represent from a policy perspective; and (ii) is 'severity' the concept that best serves that intention?

Addressing the first question of what severity is intended to represent in policy, the two versions of severity in the Norwegian context—one for the group level and one for the clinic—are intended to aid decisionmakers at different levels to prioritise in line with national guidelines, in a manner that is appropriate to their setting. Comparison of the findings in this thesis with these two versions, however, demonstrates that some of the issues members of the public consider relevant to severity, such as the effect on next-of-kin and productivity, are present only in one of them. The clinical version of severity provides a more qualitative, practical application of the term, as policy documents outline as its purpose. However, it also contains *different things* than the group-level version. By suggesting that severity, among other things, may include the effects on next-of-kin and productivity, it introduces facets of severity that are explicitly excluded from the group-level operationalisation. While severity is operationalised at the group level as absolute QALY shortfall, centred on the patient(s) and effects within the healthcare system, the clinical version widens the perspective and asks clinicians to consider broader concerns. There is thus a discrepancy between the group-level and clinical-level operationalisation of severity in Norwegian policy documents. This raises the question of whether these additional concerns, considered relevant both by policymakers who have imbedded them in the clinical operationalisation of severity, and by members of the public, are applied in any systematic way in priority-setting decisions.

Systematic resource distribution requires that those responsible for allocation are aware of their role in priority setting and of the guidelines intended to inform their decisions. Several studies demonstrate that healthcare personnel in the Norwegian context have limited awareness of the priority-setting criteria intended to guide resource allocation (Mesel, 2009; Nortvedt et al., 2008; The Directorate of Health, 2018). This is problematic: if the two operationalisations of severity emphasise different concerns, and those intended to prioritise according to one of those operationalisations have not been made sufficiently aware of their role as resource allocators, or of the guidelines intended to regulate their decisions, this raises questions of how transparent and systematic the priority-setting framework is in practice.

Furthermore, while priority setting at the clinical level is intended to provide healthcare personnel with the opportunity to consider the patient in front of them and their specific situation, they are also entrusted with the role of gatekeeper to public healthcare. Healthcare personnel must distribute resources according to both what is medically sound for the individual patient and in a manner that is in line with fair allocation of healthcare resources at the group level (Opsahl, 2020; The Norwegian Medical Association, 2021). These different roles can lead to a dilemma between professional ethics and sound priority setting, and between bedside versus group-level rationing. This dilemma speaks to the broader issue of proximity ethics, and whether healthcare personnel’s allegiance ought to be with large patient groups or with the patient facing them (Magelssen, 2018). According to proximity ethics, the clinician—as the advocate of their patient—has a moral responsibility to promote the interests of that patient, over those spatiotemporally removed from them (Magelssen, 2018; Nordhaug & Nortvedt, 2011). While the matter of proximity ethics and the role of clinicians in broader resource allocation is a debated one (Wyller, 2014, 2017), it is a heavy burden—and, as a physician myself, I would argue a big ask—for clinicians to be the sole guardians and implementors of societal concerns such as productivity in priority-setting processes.

One might question whether allowing for and encouraging integration of different concerns in decision-making at the clinical level, and omitting these at the group level, leads to a lack of transparency in how priority-setting decisions are reached—especially when the implementation of these additional concerns relies entirely on the (unknown) emphasis they are given by healthcare personnel. Relying on "clinicians' discretion" to assess the severity of individual cases (Ministry of Health and Care Services, 2016) plausibly induces inconsistencies in how priority-setting decisions are reached. If these concerns are considered relevant for the individual patient in their interaction with their physician, one could argue they ought to be considered relevant for all patients in their interaction with the healthcare system.

The second question (posed on page 8) was whether the concept of severity best serves the intentions behind the criterion. The purpose of the severity criterion is to ensure prioritisation of the severely ill. However, considering the ambiguity, polyvalence, and plurality uncovered surrounding severity, it appears that operationalising a concern for the severely ill raises more questions than it answers: who are the more severely ill? According to which parameters ought one to measure severity? From which perspective does one determine severity? Is severity an individual question or an objectively definable matter? Operationalising a concept that appears to be as complex as severity is a considerable challenge for policymakers. A worthwhile consideration, especially ahead of the 2024 white paper on priority setting, is whether this complexity is inherent to severity as a concept itself, or whether this complexity (also) stems from ambivalence in exactly what severity is intended to represent in a policy setting. Severity intuitively appears to be an important concern, and the findings presented here suggest it is associated with a variety of meanings; all of which may be relevant to priority setting.

This apparent polyvalence may be a strength of severity, because its many meanings combined appear to cover multiple different concerns we considered to be important. However, it may also represent a critical weakness: if severity is so broad and can have so many meanings, is it helpful in ensuring transparent, systematic priority setting? Because priority setting is a contentious issue, and because priority-setting outcomes are consequential and therefore important to people, transparency in how these outcomes are reached is critical. If severity is so complex that it requires two (arguably disparate) operationalisations across different levels of the healthcare sector, I would argue there is need for more conceptual work on what severity is intended to represent as a priority-setting principle, as well as a consideration of how this intention and its subsequent operationalisation compares to public views of severity.

Finally, it is worth noting that Norway is a leading country in the development of priority-setting frameworks (Wester & Bringedal, 2018). Norwegian policymakers therefore arguably have a responsibility to ensure a transparent, legitimate system that other countries can build on. Our findings suggest we still have work to do.

6.7 Methodological considerations

Strengths and limitations were discussed in each of the individual studies. This section contains a broader discussion of key methodological consideration of this thesis.

6.7.1 Design

Severity appears to be an ambiguous term and identifying and understanding public views on what severity means therefore requires an explorative approach. This idea was ingrained in the design of the three studies. The overall research question of what severity means to the general public was therefore approached from different angles: qualitatively via a thematic analysis in Study I, with the mixed-method Q methodology in Study II, and quantitatively using survey methods in Study III.

There are challenges to applying a mixed-methods approach. The differences between qualitative and quantitative approaches to research are often described as grounded in two different paradigms, with qualitative research approaching the research subject as inherently subjective and based on interpretation, and quantitative research presupposing a more singular truth (Paley, 2000). Epistemologically, qualitative research assumes that the researcher and the subject are interactively linked, whilst they are seen as independent in the quantitative paradigm (Coast, 2017; Sale et al., 2002). This risks considerable methodological tension in mixed-methods research. While mixed-methods research requires that the researcher gains a genuine understanding of contrasting methods, I believe that exploring severity from different angles, with different methodological approaches and three different data sets, is a central strength of the PhD project.

Applying thematic analysis, mixed-methods Q methodology, and quantitative survey techniques to explore the meaning of severity allowed for a broad and in-depth exploration of the concept. This represents a form of method triangulation, using multiple datasets and applying more than one approach to gain a comprehensive overview of the subject matter (Heale & Forbes, 2013). While Study III depended on the results of Study II, both studies were conducted independently from Study I. Exploring severity via method triangulation provided a rich, in-depth understanding of public views on severity, as well as an exploration of severity in a large sample representative of the Norwegian population as a whole.

There are various strengths and weaknesses to the methods applied in the three studies of this thesis. Group interviews, conducted in the format of open discussion, allow participants to lead, to introduce new topics, and the researcher must then find a way to pull these various accounts together systematically and thematically. This format also demands a lot of participants, as some of the topics may concern issues they have not previously considered, and they may therefore struggle to organise their thoughts and convey their views accurately. Q sort exercises offer a more structured, less free way to articulate one’s views, by being provided with a set of statements and ranking these. Analysis of these data is typically also a collaborative and less time-consuming exercise than e.g., thematic analysis. Finally, survey studies are even more manageable for respondents, who have a less time-consuming and perhaps simpler task—yet, the researcher is not present for data collection, and cannot help or guide respondents along the way. And both Q sort exercises and surveys, despite asking respondents to provide additional information, are less free in format, and thus limit the respondent’s opportunity to freely express their own views independently of the tasks they are given.

With substantial and varied data collection, the three studies complement each other and provide new knowledge, contributing to a deepened understanding of what severity means to the public. This mixed-methods approach also makes our results useful to policymakers. While this has been a demanding and time-intensive undertaking, I believe the mixed-methods approach of this thesis is a well-suited approach to studying a concept as complex as severity.

There are specific challenges in the progression from a Q study to a vignette-based survey. Study III builds on Study II, converting the viewpoints into vignettes. However, the findings of the survey pertain to the vignettes, rather than the original viewpoints. Respondents were asked to consider the shorter vignettes, not the full factor analysis and the resulting four viewpoints. The vignettes are adaptions of the viewpoints identified in the Q study. While the viewpoints in the Q study represent rich descriptions of severity, the vignettes compress these into 100-word summaries, and the vignettes therefore contain less detail than the viewpoints they represent. Some of the meaning and detail of the viewpoints may have been lost in the conversion into vignettes. This means that survey respondents may have responded differently to the full viewpoints in the Q study than they did to their vignette-representations However, as Q studies are difficult to perform with large samples, considering the time-intensiveness of Q sort exercises, bringing the viewpoints identified in a Q study into a survey suitable for distribution across a large sample provides a way to quantify the findings in a manner Q sort studies cannot. While there are limitations to this method, it represents our chosen approach to exploring public views (and the representation of these views) on a conceptually complex concept. While there is an increasing number of studies using Q-to-survey methods, the number of Q-to-survey studies is still limited, and even fewer adopting the vignette approach. Accordingly, Study III also makes an important contribution to the Q-to-survey methods literature, and to Q methodology overall.

Another challenge in mixed-methods research is how to explain similar results when qualitative and quantitative approaches supposedly consider different phenomena. I have spent time reflecting on this, as there are similarities in the findings of Study I and Study II (the latter of which has quantitative elements), specifically regarding the identification of subjective versus objective conceptualisations and viewpoints on severity. One might question whether the fact that some comparable findings emerged in Study I and II could be due to my own involvement in design and analysis, and assumptions and interpretations applied in the analytical process. I will return to specific steps I took to counter potential bias and presuppositions in analysis. However, while both Study I and Study II use data from the group interviews (albeit in different ways, with the former applying thematic analysis of the transcripts and the latter collecting only statements on severity), the elements of subjective and objective stood out very clearly in the analyses for the two studies—despite being conducted more than a year apart, and despite the fact that the analysis for Study II was based not on the transcripts, but on how a new group of participants ranked statements from the transcripts. Collecting statements for Study II was also a collaborative effort between myself and the PhD candidate on WP3, and these were thereby not solely selected by me. The analyses for the two studies were also overseen by two separate groups of authors, with only me and DGTW being co-authors on both studies. I believe the emergence of these elements in both studies, despite the steps taken to counter undue research bias, buttresses the validity of our findings, and attests to the existence of these conceptualisations and viewpoints on severity among members of the public.

The various steps of study design and data collection were a collaborative effort. All co-authors and SEVPRI team members made important contributions to design, and data collection was a collaborative effort between me, the Principal Investigator, and the PhD candidate on WP3. It was nonetheless a task that was first and foremost my responsibility. I spent a considerable amount of time recruiting participants for the conversations and Q sort exercises; I facilitated the majority of the conversations and Q sort exercises; and I transcribed the majority of the conversation audio recordings. By the time data collection was completed and analysis could commence, I was highly familiar with the data. Being present for the conversations and subsequently transcribing them allowed me to repeatedly and actively engage with the data, and to truly immerse myself in them. I believe this is a central strength of this thesis. The familiarity with the data and central role in data collection gave me a unique perspective on participants' views, the context in which they were shared, their body language and other non-verbal communication when sharing them, and the group setting surrounding them and the reactions of the groups to different statements. This was especially valuable to me when conducting the thematic analysis for Study I, which benefitted from a close familiarity with the primary source of the data and the possibility for a deeper understanding of the views participants had shared.

Being so familiar with the data also required a conscious effort to identify potential biases and to avoid bringing these into the analyses. If such biases are not identified, they can pose a challenge to the quality of the analyses and discussions, and familiarity with the data can become a limitation rather than a strength. Because I was so actively involved with both data collection and analysis throughout the different studies, I was diligent about identifying my own biases. The research team also made considerable efforts to ensure we discussed and identified biases (and potential biases) to avoid bringing these into our deliberations, analyses, and writing. I return to these efforts (as a team and as an individual) in the 'Reflexivity and bias' section below.

In Study I and Study II, participants were asked to fill out a questionnaire (see Chapter 4.1.1 for further details) that included an EQ-5D-5L survey (EuroQol Research Foundation, 2023).

This is a self-assessment of current health status and involves questions about mental health. Because both I and another member of the research team are healthcare personnel, with an obligation to intervene when there is a potential danger to life and health, we made sure to read the EQ-5D-5L responses shortly after the conclusion of each data collection session. In the EQ-5D-5L questions on anxiety and depression, several participants agreed with the statement: "I am extremely anxious or depressed" (EuroQol Research Foundation, 2023). In these cases, I or the other healthcare personnel team member would seek out the participant either in private before they left the premises or via telephone, seeking to ensure that they had the help they needed and that their life was not in immediate danger, according to standard guidelines. We maintained the same practice when participants shared views during the data collection sessions suggesting they had serious mental health problems.

Conducting group interviews during and after SARS-CoV-2 pandemic restrictions

I have reflected on the role of the SARS-CoV-2 pandemic and the associated restrictions regarding online recruitment and interviewing. We conducted group interviews during Spring 2021, a year after the pandemic's first wave. Following lockdown, an unprecedented share of the population was forced to work, study and socialise from home. The digital literacy of the population likely increased during this time, as digital communication tools like *Zoom* and *Microsoft Teams* became considerably more common in workplaces and universities.

It is not unlikely that recruiting for and conducting interviews following the sudden increase in the use of online meeting tools made a larger portion of the population more able and willing to participate in an online study. Some participants also stated that they were happy to participate in a study because they had more time on their hands (due to pandemic restrictions). The context of the pandemic may thus have been helpful to recruitment.

The option to recruit online and conduct group interviews online was also helpful in trying to counter the delays imposed by pandemic restrictions on physical gatherings, which delayed both the conversations and the Q sort exercises. Pandemic-related restrictions were what motivated the switch to online conversations. Because we had decided to conduct the next step of data collection, the Q sort exercises, in physical sessions, I made a considerable effort to have every step ready to begin recruiting and conducting the exercises as soon as restrictions were lifted. As a result, recruitment could begin the day after restrictions were lifted, and—as I was eager to finish data collection before potential new lockdowns—data collection was completed within seven weeks. This responsiveness would have been harder were it not for the technical tools available for online recruitment.

A final note on the effect of the SARS-CoV-2 pandemic on data collection pertains to the topic of priority setting itself. During the pandemic healthcare priority setting was at the forefront of the national and international discourse. Whether discussing the rationing of scarce vaccines, the admittance of different patient groups to hospital, or the cost and effect of different restrictions, the population became more engaged with and aware of the scarcity of certain health resources, and the need for rigorous priority setting in healthcare. It is likely that this awareness had an impact on participants' views on priority setting, their familiarity with priority-setting dilemmas, and their willingness to accept the need for priority setting at all. While we tried to elicit views relating to the meaning of severity itself, rather than severity as a priority-setting principle, the public discourse on priority setting during the pandemic may have influenced participants' views on what severe illness is.

6.7.2 Reflexivity and bias

Personal reflexivity

When analysing data, it is unavoidable that the researcher to some extent is coloured by their own preconceptions, views, and expectations. In any research, author reflexivity is therefore essential, and perhaps especially so in qualitative work (Cohen & Crabtree, 2006; Flick, 2013; Malterud, 2017). With a mixed-methods study, and especially considering the qualitative aspects of Study I and Study II, reflexivity was a central element in the planning stages, in analysis, and in discussions of findings. I sought a reflexive approach to the data throughout all three studies.

Being a physician, I made efforts to identify and be aware of my own perceptions and preconceptions regarding severity, which were likely to be influenced by my medical training and professional experiences in the healthcare sector and views on illness severity I had gathered in clinical work. While clinical experience felt valuable to my understanding of the context of this study, it also required a continuous practice of reflexivity where I sought to evaluate and critique how my subjective views and preconceptions could influence the research. I also strived to be an active listener and keep an open mind when encountering different viewpoints to my own.

There were several practices I adopted throughout the design process, data collection, analysis, and writing of the studies. A central step in my personal reflexivity involved taking time to consider, ahead of data collection, what illness severity meant to me, what I considered not to be severe, and to bring this awareness into the data collection for different stages of this thesis. This awareness helped me to prevent giving responses that aligned or collided with my own views favourable or unfavourable treatment. I was cognisant of these forms of interviewer bias (Salazar, 1990) when facilitating the conversations and Q sort exercises, and sought to bring this awareness with me to avoid asking questions or reacting to responses in a manner influenced by my own views. The insights I gained from personal reflexivity throughout the research process were important when I moved on to analysis of the Q sort exercises, analysing the factors that emerged and formulating these into factor descriptions. I also spent time considering which factor(s) I identified with, and ensured team members participated in the development of factor descriptions, as well as in development of the vignettes for the Q-to-survey study.

Team reflexivity

Throughout recruitment, conducting the conversations and Q sorts, and analysis, several decisions were made by the research team. These include choice of recruitment strategies, choice of location for conducting interviews and Q sorts, analytical approaches to the thematic analysis, selection of statements for the Q sort study, interpretations of the factors in the Q sort study, design of the survey study, exclusion criteria for the survey, vignette alignment/membership criteria for the survey data, etc. To avoid interpretive biases, this demanded a high level of reflexivity for all co-authors and team members. We sought to identify our biases via discussion and deliberation (Guest et al., 2012), and to as great an extent as possible attempt to avoid these affecting results. We strove to maintain a reflexive, collaborative approach throughout all steps of the studies.

I facilitated the majority of the group interviews, but almost every interview was conducted with another SEVPRI team member present, either as observer or alternate facilitator. This helped to avoid blind spots in the interviewing process, with another team member available to catch onto topics brought up by participants and ask follow-up questions that the facilitator might miss. For Study I, I performed the coding and analysis of the data. I made sure to seek the perspectives of co-authors on the approach I was taking before beginning coding and analysis. Two co-authors (JR and HL) read and coded three of the transcripts separately from me, and the three of us subsequently compared our coding. We were satisfied that we had approached coding in the same manner, and in line with accepted conventions for thematic analysis (Crabtree, 1999). The coding framework was based on the codes identified in the analytical process, and subsequently expanded based on the same coding principles as I continued to code the remaining transcripts. The analysis of the codework was conducted by me, based on input from all co-authors.

In Study II, the search for relevant statements in the transcripts was conducted by me and two co-authors (BJ and MB), and all transcripts were coded by me and BJ. Following the identification of relevant statements, I collected all statements in one document. These statements were organised into a coding framework based on the topic guide from the interview guide and supplemented with new codes identified when analysing the transcripts. Once all statements were organised according to the coding framework, BJ and I selected statements, designing the Q set by discussion and consensus.

I translated the final Q set to English so that co-authors RB and DGTW could review and comment on the final Q set. Once the Q sort exercises were complete and data were uploaded into the Q methodology software, all co-authors met for a physical workshop (with DGTW participating online from Canada) to analyse and discuss different factor solutions emerging from the analysis. The different factor solutions were discussed among the team, and we settled on the final factor solution by consensus. We also performed a preliminary interpretation of the factors collaboratively before I continued with a detailed analysis. The final results were discussed with all co-authors.

For Study III, a workshop was organised with all co-authors (with myself participating online, as I was in the USA for a research exchange at the time) to discuss survey design. We were joined by two experts in Q methodology (Professor Job van Exel and Professor Helen Mason), who have conducted multiple Q-to-survey studies and who provided valuable input and reflections on study design. We also held several meetings with all co-authors once data collection was completed ahead of commencing analysis, to discuss exclusion criteria and our approach to factor membership. The continued discussions, deliberations, and group effort-approach to design and analysis throughout the three studies is likely to have mitigated the risk of undue bias.

Information bias

Information bias can arise when there are systematic differences in the collection, recording or handling of data (Delgado-Rodríguez & Llorca, 2004). To avoid this, the introductory text was written to ensure that all participants were given the same information ahead of the conversations and the Q sort exercises, and we provided a similar introductory text to respondents in the survey. To avoid providing information that could influence results, this text was kept as succinct and short as possible, and the facilitators sought to avoid asking leading questions during the conversations. Having two SEVPRI team members present in the conversations also served to monitor the possible introduction of bias, and provided the opportunity to debrief after each conversation and improve facilitation skills. Furthermore, as there were several project collaborators involved in transcription of the conversations, we also made efforts to ensure we were transcribing in the same manner. Ahead of the transcription process we agreed that the transcriptions would be transcribed verbatim. We used the same coding system for all group interviews and participants, giving each interview and participant a unique, anonymous code. At the outset of the transcription phase, we also compared the first transcriptions to ensure similar techniques.

Selection bias

Recruitment for Study I and Study II was done via SEVPRI's social media platforms, via posters hung on lampposts and in shops in Oslo, and at several general practitioner offices in Bergen. While recruiting both online and in-person helped reach a greater number (and perhaps a wider array) of individuals, these methods of recruitment also impose selection biases. First, although online recruitment is efficient and enables researchers to reach more individuals, it may also lead to the exclusion of participants who are less familiar with the internet and online communication tools. That said, that some of the conversations were conducted online from whichever location participants preferred allowed participants that might otherwise have had difficulties participating to take part. It may also have felt less daunting—joining the meeting from home gave participants the option to leave whenever they wanted to. Withdrawing consent or participation in-person is arguably more socially difficult than logging off an online meeting. It also bears mentioning that, despite informing participants that they could pull out from the conversations at any point, no participants availed of this. This is encouraging, and suggests the interviews were conducted in a manner that made participants feel comfortable participating, and that the topic was engaging enough for them to remain in the interview setting from one to three hours.

The Q sort exercises were conducted in-person only, as they were intended to in the original SEVPRI study design. Due to pandemic restrictions we considered conducting them online. However, online Q sorting is still less common than in-person exercises, and online Q sorting tools are less validated than in-person participation and not necessarily as easy to conduct or complete. I tested an online Q sorting platform ahead of recruitment, and despite this solution allowing for broader recruitment across different geographical locations, I found the task of completing a Q sort exercise considerably more difficult online than in-person. In my experience, the screen was too small to provide an overview of the Q sort grid, the cards could only be read one-by-one and not placed out for comparison and simultaneous consideration, and reordering the cards once they were placed was cumbersome. I therefore made the decision to wait until pandemic restrictions were lifted to conduct the exercises in-person. Because the exercise demanded the placement of 56 individual cards across a Q sort grid, being able to move the cards physically around and see their placement clearly is likely to have had a positive effect on the degree to which participants engaged with the statements and reflected on their placement. Participants often sat with cards in their hands for a long time, placing them once before moving them elsewhere. As this was much more complicated in online platforms, conducting the Q sort exercises online could have negatively impacted the results, making participants disinclined to spend time shuffling the cards around if it was difficult to do, or choosing to exit the session if it was frustrating to complete. However, physical gatherings in two cities restricted the recruitment of participants in other areas of Norway. While we did not have the resources to travel across Norway once again for the Q sort exercises, restricting data collection to two cities may have led to the exclusion of some views. That said, results from Study III indicate that the vast majority of respondents found vignettes that resonate with their views.

Social desirability bias

Conducting the conversations in a group setting allowed participants to actively engage with each other and the different opinions that were expressed. The group setting appeared to help spur memories and thoughts for the participants, providing a wider sample of views on severity than if the interviews were conducted one-on-one. While conducting interviews in groups opened for lively discussions and an exchange of ideas and views, the group setting can also risk imposing social desirability bias (Bergen & Labonté, 2020). Participants could be inclined, consciously or unconsciously, to present views that align with what they perceived as socially acceptable within the group setting. This is particularly relevant to this thesis where participants discussed health and illness, which are sensitive topics to many. Furthermore, participants were strangers and unfamiliar with each other's health status. This may have made them hesitant to share controversial opinions or strong views, wary of offending other participants by suggesting certain conditions were or were not severe.

We attempted to circumvent social desirability bias by emphasising in the introduction of the conversations, the Q sort exercises, and the survey that there are no right or wrong answers to the question of what makes an illness severe, and that this is a topic professionals and academics still debate and disagree upon. We also emphasised that all views were welcome, and that we were seeking to identify a variety of views on severity. This concern was central to the way conversations were facilitated: we attempted to maintain a relaxed, friendly atmosphere, encouraging participants to share their views, and repeatedly pointing out that all views and opinions were welcome. When participants strongly disagreed, we emphasised that all views were relevant and interesting, and made an effort to ensure every participant was given the opportunity to share their views. While social desirability bias is difficult to avoid in qualitative research, these strategies are likely to have reduced it.

6.7.3 Exclusion of participants

In Study III, we excluded certain respondents from the analytical sample based on the time spent completing the survey, which was determined in two ways: (i) by timing how long it took members of the research team to read through all the text in the survey and complete the exercises within it; and (ii) by looking at the number of inconsistencies in respondents' vignette ranking and scoring and time spent completing the survey. These approaches led to the decision that respondents who had completed the survey in or under 3.5 minutes should be excluded from the analysis. It is possible that some of the excluded responses were valid (i.e., the responses accurately represented an excluded respondent's views). However, the proportion of excluded respondents was relatively small (6.8% of 1174 respondents) and any impact on the results is likely to be negligible.

6.7.4 Validity and generalisability

Data for the Q-to-survey study in Study III were collected from a representative sample across Norway, spanning different geographical locations and sociodemographic profiles, which increases the external validity of the study. Further, the size of the sample was relatively large (n=1094 following exclusion according to the abovementioned criterion). This is likely to have provided more precise estimates than a smaller sample would have allowed, and therefore increases my confidence in the strength of the different associations and memberships described.

An important issue in research is the degree to which findings are generalisable to a wider population, and—especially for qualitative research—what the transferability of findings are beyond the setting of that particular study (Nassaji, 2020). The generalisability, or transferability, of any study is impacted by who the participants in the study are and how they are selected (Delgado-Rodríguez & Llorca, 2004). I sought to avoid selection bias, which would negatively impact the generalisability of our findings, throughout all stages of recruitment, as outlined above. To promote generalisability, variation in the sample is key, and several steps were taken to secure variation throughout the sampling processes. We included participants from multiple different backgrounds and geographical locations in both the conversations and the Q sort exercises. However, the data in Study I and Study II are based on non-representative samples of the Norwegian population. Although we recruited purposively for both data sets and sought to reach a wide variety of individuals with different sociodemographic backgrounds, there is an overrepresentation of certain groups (women, individuals with higher education; see individual studies for more detail). While we have gathered a variety of views on severity, these methods restrict the generalisability of the findings in Study I and Study II, and it is possible there are additional views in the Norwegian population on the meaning of severity to those identified.

External validity was a central consideration in Study III, and it was important to have a nationally representative sample. The Q-to-survey study provided an exploration of the prevalence and distribution of the views described by the vignettes within a representative sample of the Norwegian population. Recruiting via the Norstat panel helped secure representation in the sample. However, internet panels may not represent the views of the wider population, and despite quota sampling via the Norstat panel, it is possible that the views of the panel do not fully reflect the views of the subpopulations the respondents from the panel are intended to represent (Scherpenzeel, 2010).

Another consideration is the role of translation, and to what degree translated material and data can fully represent the information gathered in an original-language context. I translated material pertaining to the design and data collection processes, excerpts from the data, and findings, for the purpose of presentation in the different studies. Some of the meaning, both in the way participants were presented with information, and in their responses and contributions, may have been lost in the process of translation.

A final consideration is the temporal generalisability of our findings. Our data pertain to the 2021–2023 context and may not reflect views on severity in the future. Because healthcare provision and priority setting are dynamic and subject to constant change, societal developments and influences may shape and alter the views of the public on what constitutes severe in a healthcare context. We cannot know if our findings are equally valid in years and decades ahead. The (at the time) ongoing SARS-CoV-2 pandemic may also have influenced the views of participants and made findings more contextually bound. While priority setting in healthcare is of continued relevance to the public and holds a steady presence in the public discourse (Stenmarck et al., 2021), there is no doubt that priority setting became particularly relevant as the pandemic unfolded and healthcare priority-setting dilemmas dominated the public discourse. The term 'severity' was also invoked heavily throughout the pandemic, used to describe COVID-19 illness, vaccine shortages, different patient groups' illness, etc. These societal events are likely to have impacted public views on priority setting and the meaning of severity. Findings may have been different had the studies been conducted ahead of the pandemic, or some years into the future when the pandemic is not so vividly present in everyday lives and recollections. Interestingly, however, COVID-19 illness specifically and the pandemic generally rarely came up in the group interviews. Furthermore, a study by Arroyos Calvera and colleagues on public preferences regarding efficiency and equity in policies found that preferences pre- versus post-pandemic were remarkably stable (Arroyos Calvera et al., 2023)

I have studied the views on severity across the Norwegian population. Results may have differed if the studies were conducted in a different country and healthcare system, where other social values dominate (Littlejohns et al., 2012). Although data collection for Study I and Study II was conducted with small, non-representative samples of the Norwegian population, they present views on severity that may still be relevant in other countries. Severity is a priority-setting principle in several healthcare systems, and researchers and decisionmakers in other jurisdictions may find some of the insights useful—while applying a certain degree of caution regarding generalisability. Knowledge about societal views on a given topic which is relevant in different countries allows for cross-country comparisons (Mason et al., 2016). Knowledge about public views on severity in Norway can enable exchange of experiences across different countries applying severity as a priority-setting principle, and provide valuable knowledge for policymakers. I am confident our studies will make a positive contribution in this regard.

6.7.5 Intuitions versus considered judgements

Whether, to what degree, and how public views are to be included in policy-formulation is a much-debated issue. Several authors have argued that public views ought to be included in reflective equilibrium processes, either as intuitions or as considered judgments (Baderin, 2017; Gustavsson & Lindblom, 2023; Savulescu et al., 2021). I recently co-authored a paper on publicly informed reflective equilibrium processes based on the findings of Study II (Jølstad et al., *forthcoming*). This led me to reflect on the nature of the data collected throughout the different stages of this PhD project, and whether the views gathered represent intuitions or considered judgements.

In the conversations and the Q sort exercises, participants sat for one to three hours and deliberated, either together (conversations) or alone (Q sort exercises), on what severity meant to them. They had time to consider their views and play out different scenarios, to change their mind if they were convinced by other participants' arguments, or to champion their views as they developed their own arguments in defence of them. As such, these data points arguably represent considered judgements. In the survey, however, average completion time was 11 minutes. Considering the length of the survey and the time required to fill it out, the average respondent cannot have spent long considering the different vignettes. The responses in the survey are therefore more likely to resemble intuitions, or gut reactions. The data collection method, whether conversations, Q sort exercises, or survey, may have been instrumental in determining the nature of the views on severity as either intuitions or considered judgments. It could be of interest for future research to consider whether the results of the Q-to-survey study might have been different had respondents participated in a conversation on severity beforehand. If so, their responses could to a greater degree be said to represent considered judgments. It could also be interesting to explore whether responses differ if the survey is retaken at a later point in time, when respondents would have had time to reflect on the dilemmas and questions the survey presented.

7. Conclusions

This thesis brings together knowledge on public views on severity and the distribution of these views across the Norwegian population, making several contributions. I have condensed these contributions into the following three points:

- Study I and Study II provide detailed descriptions of how severity is viewed and conceptualised by members of the public. The findings in these studies demonstrate that, while QALY shortfall outcomes appear to be considered relevant for understandings of severity, the citizenry have additional concerns which are not accounted for by QALY shortfall operationalisations of severity.

- Study III provides a description of the distribution, across the Norwegian population, of the viewpoints on severity identified in Study II. The results demonstrate that there is support for all summary descriptions of the viewpoints (i.e., the vignettes), and no one vignette stands out as representative of the entire population. These are also distributed evenly across the population.

- This thesis also makes methodological contributions. First, to Q methodology by contributing to the relatively new research field of Q-to-survey studies and pushing the frontier of vignette-based surveys. Second, it demonstrates the value of applying a mixed-methods approach when exploring complex concepts. I believe the combination of qualitative and quantitative approaches across the studies in this thesis represents a well-suited approach to exploring phenomenological complexity, while also providing policymakers with useful knowledge which can be applied in a policy setting.

Based on the findings presented in this thesis, I draw the following conclusions. First, the different conceptualisations and viewpoints on severity presented in Study I and Study II suggest that there is a plurality of views on severity which are not necessarily easily reconcilable. They represent different approaches to severity, some with overlap between them but others vastly different. Furthermore, the findings in Study III demonstrate that the identified viewpoints (represented by vignettes) have support in the wider population. *In toto*, these findings suggest that severity is a polyvalent concept. I argue that this polyvalence can in-part be explained by (i) the conceptual complexity of severity, containing both descriptive and normative elements; and (ii) three different qualities inherent to severity, representing a framework for public views on severity consisting of disease severity, illness severity, and sickness severity. Approaching severity through this framework provides a more wholesome understanding of public views on severity.

A second conclusion is that there appears to be a misalignment between public views and policy operationalisations of severity. Current operationalisations fail to capture many of the public's common intuitions about severity. This misalignment poses a challenge to policymakers and has potentially detrimental effects for priority-setting work. At best, it is likely to lead to miscommunications between policymakers and the public and increase tensions surrounding priority setting in healthcare. At worst, if accompanied by a lack of policy responsiveness to knowledge about public views, this misalignment can jeopardise the democratic legitimacy of priority-setting frameworks.

Finally, I argue that, to avoid misalignment, public views ought to be incorporated at early stages of policy development. While Norway has a long tradition of working systematically with priority-setting frameworks, and while policymakers appear to be aware of the complexity inherent to severity, actively seeking out and exploring public views on the meaning of severity has not been an established part of that tradition. It remains to be seen if, once provided with evidence about current misalignments, there is political will to address this misalignment, and whether future iterations of priority-setting principles will involve more active exploration of and engagement with the views of the public.

References

Abelson, J., Blacksher, E. A., Li, K. K., Boesveld, S. E., & Goold, S. D. (2013). Public Deliberation in Health Policy and Bioethics: Mapping an emerging, interdisciplinary field. *Journal of Deliberative Democracy*, *9*(1), Article 1. https://doi.org/10.16997/jdd.157

Al-Janabi, H., van Exel, J., Brouwer, W., & Coast, J. (2016). A Framework for Including Family Health Spillovers in Economic Evaluation. *Medical Decision Making*, *36*(2), 176–186. https://doi.org/10.1177/0272989X15605094

Alvesson, M., & Sköldberg, K. (2017). *Reflexive Methodology: New Vistas for Qualitative Research*. SAGE.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. https://doi.org/10.1176/appi.books.9780890425596

Arroyos Calvera, D., Covey, J., & McDonald, R. (2023). Are distributional preferences for safety stable?: A longitudinal analysis before and after the COVID-19 outbreak. *Social Science and Medicine*, *324*(115855). https://doi.org/10.1016/j.socscimed.2023.115855

Baderin, A. (2017). Reflective Equilibrium: Individual or Public? *Social Theory and Practice*, *43*(1), 1–28.

Baker, R., Bateman, I., Donaldson, C., Jones-Lee, M., Lancsar, E., Loomes, G., Mason, H., Odejar, M., Prades, J.-L. P., & Robinson, A. (2010). Weighting and valuing quality-adjusted life-years using stated preference methods: Preliminary results from the Social Value of a QALY Project. *Health Technology Assessment (Winchester, England)*, *14*(27), 1-+.

Baker, R., Mason, H., McHugh, N., & Donaldson, C. (2021). Public values and plurality in health priority setting: What to do when people disagree and why we should care about reasons as well as choices. *Social Science & Medicine*, *277*, 113892. https://doi.org/10.1016/j.socscimed.2021.113892

Baker, R., van Exel, Job, Mason, Helen, & Stricklin, Michael. (2010). Connecting Q & Surveys: Three Methods to Explore Factor Membership in Large Samples. *Operant Subjectivity: The International Journal of Q Methodology*, *34(1)*, 38–58.

Baker, R., Wildman, J., Mason, H., & Donaldson, C. (2014). Q-Ing for Health—A New Approach to Eliciting the Public's Views on Health Care Resource Allocation. *Health Economics*, *23*(3), 283–297. https://doi.org/10.1002/hec.2914

Baker, Rachel, McHugh, Neil, & Mason, Helen. (2017). Constructing statement sets for use in Q methodology studies. In J. Coast (Ed.), *Qualitative Methods for Health Economics*. Rowman & Littlefield.

Baltussen, R., Jansen, M. P. M., Bijlmakers, L., Grutters, J., Kluytmans, A., Reuzel, R. P., Tummers, M., & der Wilt, G. J. van. (2017). Value Assessment Frameworks for HTA Agencies: The Organization of Evidence-Informed Deliberative Processes. *Value in Health*, *20*(2), 256–260. https://doi.org/10.1016/j.jval.2016.11.019

Banasick, S. (2019). KADE: A desktop application for Q methodology. *Journal of Open Source Software*, *4*(36), 1360. https://doi.org/10.21105/joss.01360

Barbu, C. M. (2014). *Zoom: A spatial data visualization tool* (2.0.6) [Computer software].

Barra, M., Broqvist, M., Gustavsson, E., Henriksson, M., Juth, N., Sandman, L., & Solberg, C. T. (2020). Severity as a Priority Setting Criterion: Setting a Challenging Research Agenda. *Health Care Analysis*, *28*(1), 25–44. https://doi.org/10.1007/s10728-019-00371-z

Bergen, N., & Labonté, R. (2020). 'Everything Is Perfect, and We Have No Problems': Detecting and Limiting Social Desirability Bias in Qualitative Research. *Qualitative Health Research*, *30*(5), 783–792. https://doi.org/10.1177/1049732319889354

Bjørnelv, G. W., & Melberg, H. O. (2023). Penger, prinsipper og prioritering. *Tidsskrift for Den norske legeforening*. https://doi.org/10.4045/tidsskr.23.0210

Blankholm, A., Sandvik, E., Løseth, B., Magnussen, J., Kjellevold, A., Schou, P., Mosti, I. L., Hembre, U., Kreso, A., Sellæg, W. F., Galaaen, V., Haver, H., Berg-Jensen, J., Haugland, K. K., Åsholt, Y., & Brelin, P. (2018). *Det viktigste først—Prinsipper for prioritering i den kommunale helse- og omsorgstjenesten og for offentlig finansiert tannhelsetjenester* (Norges offentlige utredninger, p. 148) [NOU]. Helse- og Omsorgsdepartementet.

Boardman, F. K., & Clark, C. C. (2022). What is a 'serious' genetic condition? The perceptions of people living with genetic conditions. *European Journal of Human Genetics*, *30*(2), 160–169. https://doi.org/10.1038/s41431-021-00962-2

Bognar, G. (2012). Empirical and Armchair Ethics. *Utilitas*, *24*(4), 467–482. https://doi.org/10.1017/S0953820812000027

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brock, D. (2002a). Priority to the worse off in healthcare resource prioritisation. In R. Rhodes, M. P. Battin, & A. Silvers (Eds.), *Medicine and Social Justice: Essays on the Distribution of Health Care*. Oxford University Press.

Brock, D. (2002b). The separability of health and well-being. In C. J. L. Murray & World Health Organization (Eds.), *Summary measures of population health: Concepts, ethics, measurement, and applications*. World Health Organization.

Broqvist, M. (2018). *Asking the public: Citizens′ views on priority setting and resource allocation in democratically governed healthcare*. https://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-146217

Broqvist, M., & Garpenby, P. (2014). To accept, or not to accept, that is the question: Citizen reactions to rationing. *Health Expectations*, *17*(1), 82–92. https://doi.org/10.1111/j.1369-7625.2011.00734.x

Broqvist, M., Sandman, L., Garpenby, P., & Krevers, B. (2018). The meaning of severity – do citizenś views correspond to a severity framework based on ethical principles for priority setting? *Health Policy*, *122*(6), 630–637. https://doi.org/10.1016/j.healthpol.2018.04.005

Brown, S. R. (1993). A Primer on Q Methodology. *Operant Subjectivity*, *16*.

Brown, S. R. (2002). Q technique and questionnaires. *Operant Subjectivity*, *25*, 117–126.

Butler, D. M., & Dynes, A. M. (2016). How Politicians Discount the Opinions of Constituents with Whom They Disagree. *American Journal of Political Science*, *60*(4), 975–989. https://doi.org/10.1111/ajps.12206

Butler, J. (1999). *The Ethics of Health Care Rationing*. SAGE.

Calltorp, J. (1999). Priority setting in health policy in Sweden and a comparison with Norway. *Health Policy*, *50*(1), 1–22. https://doi.org/10.1016/S0168-8510(99)00061-5

Canaway, A., Al-Janabi, H., Kinghorn, P., Bailey, C., & Coast, J. (2019). Close-Person Spill-Overs in End-of-Life Care: Using Hierarchical Mapping to Identify Whose Outcomes to Include in Economic Evaluations. *PharmacoEconomics*, *37*(4), 573–583. https://doi.org/10.1007/s40273-019-00786-5

Centre for Medicare and Medicaid Services. (2021). *NHE Fact Sheet | Centre for Medicare and Medicaid Services*. https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nhe-fact-sheet

Clark, S., & Weale, A. (2012). Social values in health priority setting: A conceptual framework. *Journal of Health Organization and Management*, *26*(3), 293–316. https://doi.org/10.1108/14777261211238954

Cleary, P. D. (1997). Subjective and Objective Measures of Health: Which is Better When? *Journal of Health Services Research & Policy*, *2*(1), 3–4. https://doi.org/10.1177/135581969700200102

Coast, J. (Ed.). (2017). *Qualitative Methods for Health Economics*. Rowman & Littlefield International. https://www.adlibris.com/no/bok/qualitative-methods-for-health-economics-9781783485628

Cohen, J. T., & Crabtree, B. F. (2006). *Qualitative research guidelines project*. http://www.qualres.org/HomeSemi-3629.html

Commission on Social Determinants of Health. (2008). *Closing the gap in a generation: Health equity through action on the social determinants of health*. World Health Organization. https://www.who.int/initiatives/action-on-the-social-determinants-of-health-for-advancing-equity/world-report-on-social-determinants-of-health-equity/commission-on-social-determinants-of-health

Cookson, R. (2013). Can the NICE “End-of-Life Premium” Be Given a Coherent Ethical Justification? *Journal of Health Politics, Policy and Law*, *38*(6), 1129–1148. https://doi.org/10.1215/03616878-2373166

Cookson, R., McCabe, C., & Tsuchiya, A. (2008). Public healthcare resource allocation and the Rule of Rescue. *Journal of Medical Ethics*, *34*(7), 540–544. https://doi.org/10.1136/jme.2007.021790

Crabtree, B. F. (1999). *Doing Qualitative Research*. SAGE.

Culyer, A. J. (2017). Ethics, priorities and cancer. *Journal of Cancer Policy*, *11*, 6–11. https://doi.org/10.1016/j.jcpo.2016.09.007

Dahl, R. A. (2008). *Polyarchy: Participation and Opposition*. Yale University Press.

Damar, E. A., & Sali, P. (2022). *Q Methodology: A Concise Overview* [Chapter]. Methodological Innovations in Research and Academic Writing; IGI Global. https://doi.org/10.4018/978-1-7998-8283-1.ch001

Daniels, N. (2000). Accountability for reasonableness. *BMJ : British Medical Journal*, *321*(7272), 1300–1301.

Daniels, N. (2007). *Just Health: Meeting Health Needs Fairly*. Cambridge University Press; Cambridge Core. https://doi.org/10.1017/CBO9780511809514

Daniels, N., & Sabin, J. (1998). The Ethics Of Accountability In Managed Care Reform: Recent efforts at reforming managed care practices have one thing in common: a call for accountability to consumers. *Health Affairs*, *17*(5), 50–64. https://doi.org/10.1377/hlthaff.17.5.50

Daniels, N., & Sabin, J. E. (2002). *Setting Limits Fairly: Can we learn to share medical resources?* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195149364.001.0001

Delgado-Rodríguez, M., & Llorca, J. (2004). Bias. *Journal of Epidemiology & Community Health*, *58*(8), 635–641. https://doi.org/10.1136/jech.2003.008466

Diederich, A., Swait, J., & Wirsik, N. (2012). Citizen Participation in Patient Prioritization Policy Decisions: An Empirical and Experimental Study on Patients’ Characteristics. *PLOS ONE*, *7*(5), e36824. https://doi.org/10.1371/journal.pone.0036824

Dineen, C. (2011). Finding the Right Way to Ration. *The American Journal of Bioethics*, *11*(7), 26–28. https://doi.org/10.1080/15265161.2011.578193

Dive, L., Archibald, A. D., Freeman, L., & Newson, A. J. (2023). How should severity be understood in the context of reproductive genetic carrier screening? *Bioethics*, *37*(4), 359–366. https://doi.org/10.1111/bioe.13136

Dolan, P., & Shaw, R. (2003). A note on the relative importance that people attach to different factors when setting priorities in health care. *Health Expectations*, *6*(1), 53–59. https://doi.org/10.1046/j.1369-6513.2003.00210.x

Dolan, P., Shaw, R., Tsuchiya, A., & Williams, A. (2005). QALY maximisation and people's preferences: A methodological review of the literature. *Health Economics*, *14*(2), 197–208. https://doi.org/10.1002/hec.924

Donaldson, C., Baker, R., Mason, H., Pennington, M., Bell, S., Lancsar, E., Jones-Lee, M., Wildman, J., Robinson, A., Bacon, P., Olsen, J. A., Gyrd-Hansen, D., Kjær, T., Bech, M., Nielsen, J. S., Persson, U., Bergman, A., Protiére, C., Moatti, J. P., … Shackley, P. (2010). *EuroVaQ Fina Report* (Project Final Report SP5A-CT-2007–044172). https://research.ncl.ac.uk/eurovaq/EuroVaQ_Final_Publishable_Report_and_Appendices.pdf

Douven, I. (2021). Abduction. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2021/entries/abduction/

Drost, R. M. W. A., Putten, I. M. van der, Ruwaard, D., Evers, S. M. A. A., & Paulus, A. T. G. (2017). Conceptualisations of the societal perspective within economic evaluations: A systematic review. *International Journal of Technology Assessment in Health Care*, *33*(2), 251–260. https://doi.org/10.1017/S0266462317000526

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford University Press.

Esaiasson, P., & Wlezien, C. (2017). Advances in the Study of Democratic Responsiveness: An Introduction. *Comparative Political Studies*, *50*(6), 699–710. https://doi.org/10.1177/0010414016633226

EuroQol Research Foundation. (2023). *EQ-5D-5L – EQ-5D*. https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/

Exter, A. den, & Buijsen, M. A. J. M. (Eds.). (2012). *Rationing Health Care: Hard Choices and Unavoidable Trade-offs*. Maklu.

Flick, U. (2013). *The SAGE Handbook of Qualitative Data Analysis*. SAGE.

Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., & Grimshaw, J. M. (2010). What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology & Health*, *25*(10), 1229–1245. https://doi.org/10.1080/08870440903194015

Franken, M., Stolk, E., Scharringhausen, T., de Boer, A., & Koopmanschap, M. (2015). A comparative study of the role of disease severity in drug reimbursement decision making in four European countries. *Health Policy (Amsterdam, Netherlands)*, *119*(2), 195–202. https://doi.org/10.1016/j.healthpol.2014.10.007

Frankish, C. J., Kwan, B., Ratner, P. A., Wharf Higgins, J., & Larsen, C. (2002). Challenges of citizen participation in regional health authorities. *Social Science & Medicine*, *54*(10), 1471–1480. https://doi.org/10.1016/S0277-9536(01)00135-6

Friedman, Y. (2021). On recovery: Re-directing the concept by differentiation of its meanings. *Medicine, Health Care and Philosophy*, *24*(3), 389–399. https://doi.org/10.1007/s11019-021-10014-7

Friedman, Y., Watzl, S., Oftedal, G., & Gørbitz, C. H. (2022). Kurering, heling og habilitering. *Tidsskrift for Den norske legeforening*. https://doi.org/10.4045/tidsskr.21.0653

Green, C. (2009). Investigating public preferences on ‘severity of health’ as a relevant condition for setting healthcare priorities. *Social Science & Medicine*, *68*(12), 2247–2255. https://doi.org/10.1016/j.socscimed.2009.03.020

Grossman, L. A. (2013). The Origins of American Health Libertarianism. *Yale Journal of Health Policy, Law and Ethics*, *13*, 76.

Gu, Y., Lancsar, E., Ghijben, P., Butler, J. R., & Donaldson, C. (2015). Attributes and weights in health care priority setting: A systematic review of what counts and to what extent. *Social Science & Medicine*, *146*, 41–52. https://doi.org/10.1016/j.socscimed.2015.10.005

Guest, G., M.MacQueen, K., & E.Namey, E. (2012). *Applied Thematic Analysis*. SAGE Publications, Inc. https://doi.org/10.4135/9781483384436

Gustavsson, E., Juth, N., Lärfars, G., Raaschou, P., & Sandman, L. (2023). Should relational effects be considered in health care priority setting? *Bioethics*, *1*(6). https://doi.org/10.1111/bioe.13189

Gustavsson, E., & Lindblom, L. (2023). Justification of principles for healthcare priority setting: The relevance and roles of empirical studies exploring public values. *Journal of Medical Ethics*, jme-2022-108702. https://doi.org/10.1136/jme-2022-108702

Gutmann, A., & Thompson, D. F. (2004). *Why Deliberative Democracy?* Princeton University Press.

Gyrd-Hansen, D. (2004). Investigating the social value of health changes. *Journal of Health Economics*, *23*(6), 1101–1116. https://doi.org/10.1016/j.jhealeco.2004.02.002

Ham, C. (1997). Priority setting in health care: Learning from international experience. *Health Policy*, *42*(1), 49–66. https://doi.org/10.1016/S0168-8510(97)00054-7

Hannawa, A. F., Wu, A. W., Kolyada, A., Potemkina, A., & Donaldson, L. J. (2022). The aspects of healthcare quality that are important to health professionals and patients: A qualitative study. *Patient Education and Counseling*, *105*(6), 1561–1570. https://doi.org/10.1016/j.pec.2021.10.016

Hansson, L. F., Norheim, O. F., & Ruyter, K. W. (1994). Equality, Explicitness, Severity, and Rigidity: The Oregon Plan Evaluated from a Scandinavian Perspective. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, *19*(4), 343–366. https://doi.org/10.1093/jmp/19.4.343

Hausman, D. (2019a). The significance of 'severity'. *Journal of Medical Ethics*, *45*(8), 545–551. https://doi.org/10.1136/medethics-2018-105058

Hausman, D. (2019b). The significance of 'severity'. *Journal of Medical Ethics*, *45*(8), 545–551. https://doi.org/10.1136/medethics-2018-105058

Heale, R., & Forbes, D. (2013). Understanding triangulation in research. *Evidence-Based Nursing*, *16*(4), 98–98. https://doi.org/10.1136/eb-2013-101494

Hess, R. D., & Hink, D. L. (1959). A Comparison of Forced vs. Free Q-Sort Procedure. *The Journal of Educational Research*, *53*(3), 83–90.

Hirose, I., & Bognar, G. (2014). *The ethics of health care rationing: An introduction*. Routledge.

Hofmann, B. (2002). On the Triad Disease, Illness and Sickness. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, *27*(6), 651–673. https://doi.org/10.1076/jmep.27.6.651.13793

Hofmann, B. (2016). Disease, Illness, and Sickness. In M. Solomon, J. R. Simon, & H. Kincaid (Eds.), *The Routledge Companion to Philosophy of Medicine*. https://www.routledgehandbooks.com/doi/10.4324/9781315720739.ch2

Hofmann, B., & Wilkinson, S. (2016). Mange betegnelser for sykdom. *Tidsskrift for Den norske legeforening*. https://doi.org/10.4045/tidsskr.16.0316

Hofstede, G. (1984). The Cultural Relativity of the Quality of Life Concept. *Academy of Management Review*, *9*(3), 389–398. https://doi.org/10.5465/amr.1984.4279653

Horn, S. A., Barra, M., Norheim, O. F., & Solberg, C. T. (2021). Public health priority setting: A case for priority to the worse off in well-being during the COVID-19 pandemic. *Etikk i Praksis - Nordic Journal of Applied Ethics*, *2*, Article 2. https://doi.org/10.5324/eip.v15i2.4097

Horn, S. A., Jølstad, B., Barra, M., & Solberg, C. T. (2021). *Er covid-19 alvorlig? Det norske alvorlighetskriteriet for helseprioriteringer i møte med en pandemi*. Tidsskrift for Velferdsforskning. https://www.idunn.no/doi/10.18261/issn.2464-3076-2021-02-01

Høydahl, E. (2020). *Sentralitetsindeksen. Oppdatering med 2020-kommuner*. Statistics Norway. https://www.ssb.no/befolkning/artikler-og-publikasjoner/_attachment/413602?_ts=17085d29f50

Hughes, R., & Huby, M. (2004). The construction and interpretation of vignettes in social research. *Social Work and Social Sciences Review*, *11*(1), Article 1. https://doi.org/10.1921/swssr.v11i1.428

Husom, N. (2000). Prioritering – politisk likhetsideal gir medisinsk ubehag. *Tidsskrift for Den norske legeforening*. https://tidsskriftet.no/2000/10/nyheter-og-reportasjer/prioritering-politisk-likhetsideal-gir-medisinsk-ubehag

Jacobsson, F., Carstensen, J., & Borgquist, L. (2005). Caring externalities in health economic evaluation: How are they related to severity of illness? *Health Policy*, *73*(2), 172–182. https://doi.org/10.1016/j.healthpol.2004.11.008

Jedeloo, S., van Staa, A., Latour, J. M., & van Exel, N. J. A. (2010). Preferences for health care and self-management among Dutch adolescents with chronic conditions: A Q-methodological investigation. *International Journal of Nursing Studies*, *47*(5), 593–603. https://doi.org/10.1016/j.ijnurstu.2009.10.006

Kinghorn, P., & Coast, J. (2019). Appropriate frameworks for economic evaluation of end of life care: A qualitative investigation with stakeholders. *Palliative Medicine*, *33*(7), 823–831. https://doi.org/10.1177/0269216319839635

Lehoux, P., Daudelin, G., Demers-Payette, O., & Boivin, A. (2009). Fostering deliberations about health innovation: What do we want to know from publics? *Social Science & Medicine*, *68*(11), 2002–2009. https://doi.org/10.1016/j.socscimed.2009.03.017

Lim, M. K., Bae, E. Y., Choi, S.-E., Lee, E. K., & Lee, T.-J. (2012). Eliciting public preference for health-care resource allocation in South Korea. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, *15*(1 Suppl), S91-94. https://doi.org/10.1016/j.jval.2011.11.014

Lindemark, F., Norheim, O. F., & Johansson, K. A. (2014). Making use of equity sensitive QALYs: A case study on identifying the worse off across diseases. *Cost Effectiveness and Resource Allocation*, *12*(1), Article 1. https://doi.org/10.1186/1478-7547-12-16

Linley, W. G., & Hughes, D. A. (2013). Societal Views on Nice, Cancer Drugs Fund and Value-Based Pricing Criteria for Prioritising Medicines: A Cross-Sectional Survey of 4118 Adults in Great Britain. *Health Economics*, *22*(8), 948–964. https://doi.org/10.1002/hec.2872

Littlejohns, P., Weale, A., Chalkidou, K., Faden, R., & Teerawattananon, Y. (2012). Social values and health policy: A new international research programme. *Journal of Health Organization and Management*, *26*(3), 285–292. https://doi.org/10.1108/14777261211238945

Litva, A., Coast, J., Donovan, J., Eyles, J., Shepherd, M., Tacchi, J., Abelson, J., & Morgan, K. (2002). 'The public is too subjective': Public involvement at different levels of health-care decision making. *Social Science & Medicine*, *54*(12), 1825–1837. https://doi.org/10.1016/S0277-9536(01)00151-4

Lønning, I., Album, D., Benkow, J., Brofoss, K., Brundtland, G., Buttedahl, A., Evensen, S., Nesheim, B.-G., Sandmo, A., Traaholt, I., Shetelig, A. W., Wislæff, N. F., Ødegård, L., Aase, K., Halvorsen, M., Husbyn, H., & Norheim, O. F. (1997). *Prioritering på ny—Gjennomgang av retningslinjer for piroriteringer innen norsk helsetjeneste* (NOU 1997:18; Norges offentlige utredninger, p. 225). Sosial- og helsedepartementet.

Lønning, I., Asbøll, E., Berge, T. E., Evensen, Å. R., Grund, J., Gunbjørud, A. B., Holm, N., Kristiansen, E., Normann, T., Rasmussen, K., Thu, K., Tranøy, K. E., Myhre, L., & Brofoss, K. (1987). *Retningslinjer for prioriteringer innen norsk helsetjeneste* (NOU 1987:23; Norges offentlige utredninger, p. 108). Sosialdepartementet.

MacKillop, E., & Sheard, S. (2018). Quantifying life: Understanding the history of Quality-Adjusted Life-Years (QALYs). *Social Science & Medicine*, *211*, 359–366. https://doi.org/10.1016/j.socscimed.2018.07.004

Magelssen, M. (2018). Er prioritering forenlig med helsepersonells profesjonsetikk? In H. Sinding Aasen, B. Bringedal, K. Bærøe, & A.-M. Magnussen (Eds.), *Prioritering, styring og likebehandling: Utfordringer i norsk helsetjeneste*. Cappelen Damm Akademisk.

Magnussen, J., Aaserud, M., Granaas, T., Magelssen, M., Syse, A., Celius, E. G., Klovning, A., & Syversen, I. D. (2015). *På ramme alvor—Alvorlighet og prioritering* (p. 71) [Rapport fra arbeidsgruppe nedsatt av HOD]. Department of Health.

Malterud, K. (2012). *Fokusgrupper som forskningsmetode for medisin og helsefag*. Universitetsforlaget.

Malterud, K. (2017). *Kvalitative metoder i medisinsk forskning. En innføring.* (4th ed.). Universitetsforlaget.

Malterud, K. (2019). *Qualitative Metasynthesis: A Research Method for Medicine and Health Sciences* (1st edition). Routledge.

Mander, T. (2014). Longevity and healthy ageing – Will healthcare be drowned by the grey Tsunami or sunk by the demographic iceberg? *Post Reproductive Health*, *20*(1). https://journals-sagepub-com.ezproxy.uio.no/doi/full/10.1177/1754045313515208?casa_token=irrJC8jSf7wAAAAA%3A1NI9mp2xPCHFv4mBtHi16Txf-T62T5v9Ky5QVGa_FnWUGKi-sqsPo7X5VrKVf_zEYme0FgwaMYEmUQ

Marsh, K., Lanitis, T., Neasham, D., Orfanos, P., & Caro, J. (2014). Assessing the Value of Healthcare Interventions Using Multi-Criteria Decision Analysis: A Review of the Literature. *PharmacoEconomics*, *32*(4), 345–365. https://doi.org/10.1007/s40273-014-0135-0

Martin, D. K., Giacomini, M., & Singer, P. A. (2002). Fairness, accountability for reasonableness, and the views of priority setting decision-makers. *Health Policy*, *61*(3), 279–290. https://doi.org/10.1016/S0168-8510(01)00237-8

Mason, H., van Exel, J., Baker, R., Brouwer, W., Donaldson, C., Donaldson, C., Baker, R., Mason, H., Pennington, M., Bell, S., Jones-Lee, M., Wildman, J., Lancsar, E., Robinson, A., Bacon, P., Olsen, J. A., Gyrd-Hansen, D., Kjaer, T., Beck, M., … Shackley, P. (2016). From representing views to representativeness of views: Illustrating a new (Q2S) approach in the context of health care priority setting in nine European countries. *Social Science & Medicine*, *166*, 205–213. https://doi.org/10.1016/j.socscimed.2016.08.036

Maynard, A., & Bloor, K. (1998). Our Certain Fate: Rationing in Health Care. *OHE Monograph*, Article 000443. https://ideas.repec.org//p/ohe/monogr/000443.html

McHugh, N. (2022). Eliciting public values on health inequalities: Missing evidence for policy windows? *Evidence & Policy*, *18*(4), 733–745. https://doi.org/10.1332/174426421X16286783870175

McKeown, B., & Thomas, D. (2013). *Q Methodology*. SAGE Publications, Inc. https://doi.org/10.4135/9781483384412

Mechanic, D. (1997). Muddling Through Elegantly: Finding The Proper Balance In Rationing. *Health Affairs*, *16*(5), 83–92. https://doi.org/10.1377/hlthaff.16.5.83

Mesel, T. (2009). *Legers profesjonsetikk: Refleksjon og mestring i en sykehuskultur*. Høyskoleforlaget. https://www.akademika.no/medisin-helse-og-psykologi/legers-profesjonsetikk/9788276348231

Ministry of Health and Care Services. (1999). *Meld. St. 26 (1999-2000)*.

Ministry of Health and Care Services. (2020). *Meld. St. 38 (2020-2021) Nytte, ressurs og alvorlighet—Prioritering i helse- og omsorgstjenesten*. Helse- og omsorgsdepartmentet. https://www.regjeringen.no/contentassets/85eb048de42342e4b789913cb8b1b218/no/pdfs/stm202020210038000dddpdfs.pdf

Ministry of Health and Care Services. (2016). *Meld. St. 34 (2015–2016) Verdier i pasientens helsetjeneste* [Stortingsmelding]. Regjeringen.No. https://www.regjeringen.no/contentassets/439a420e01914a18b21f351143ccc6af/no/pdfs/stm201520160034000dddpdfs.pdf

Ministry of Health and Care Services. (2019, November 22). *Meld. St. 7 (2019–2020)* [Stortingsmelding]. Regjeringen.no; regjeringen.no. https://www.regjeringen.no/no/dokumenter/meld.-st.-7-20192020/id2678667/

Ministry of Health and Care Services. (2023a, February 13). *Varig styrking av sykehusene med 2,5 milliarder kroner* [Pressemelding]. Regjeringen.no; regjeringen.no. https://www.regjeringen.no/no/aktuelt/varig-styrking-av-sykehusene-med-25-milliarder-kroner/id2962939/

Ministry of Health and Care Services. (2023b, July 7). *Regjeringen setter ned tre ekspertgrupper til prioriteringsmeldingen* [Nyhet]. Regjeringen.no; regjeringen.no. https://www.regjeringen.no/no/aktuelt/regjeringen-setter-ned-tre-ekspertgrupper-til-prioriteringsmeldingen/id2989600/

Mitton, C., Smith, N., Peacock, S., Evoy, B., & Abelson, J. (2009). Public participation in health care priority setting: A scoping review. *Health Policy*, *91*(3), 219–228. https://doi.org/10.1016/j.healthpol.2009.01.005

Mobbs, R. J. (2021). From the Subjective to the Objective era of outcomes analysis: How the tools we use to measure outcomes must change to be reflective of the pathologies we treat in spinal surgery. *Journal of Spine Surgery (Hong Kong)*, *7*(3), 456–457. https://doi.org/10.21037/jss-2021-2

Morrell, L., Ii, S. S., Wordsworth, S., Wilson, R., Rees, S., & Barker, R. (2018). Cancer as the “perfect storm”? A qualitative study of public attitudes to health conditions. *Health Science Reports*, *1*(1), e16. https://doi.org/10.1002/hsr2.16

Mullen, P. M. (1999). Public involvement in health care priority setting: An overview of methods for eliciting values. *Health Expectations*, *2*(4), 222–234. https://doi.org/10.1046/j.1369-6513.1999.00062.x

Nassaji, H. (2020). Good qualitative research. *Language Teaching Research*, *24*(4), 427–431. https://doi.org/10.1177/1362168820941288

National Institute for Health and Care Excellence. (2022). *NICE health technology evaluations: The manual*. 181.

National Research Council (US), & Institute of Medicine (US). (2013). Public Health and Medical Care Systems. In S. H. Woolf & L. Aron (Eds.), *U.S. Health in International Perspective: Shorter Lives, Poorer Health*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK154484/

NEM. (2009). *Guidelines for Research Ethics and Scientific Assessment of Qualitative Research Projects in Medicine and Healthcare*. Forskningsetikk. https://www.forskningsetikk.no/retningslinjer/med-helse/vurdering-av-kvalitative-forskningsprosjekt-innen-medisin-og-helsefag/

NESH. (2021). *Guidelines for Research Ethics in the Social Sciences and the Humanities*. Forskningsetikk. https://www.forskningsetikk.no/en/guidelines/social-sciences-humanities-law-and-theology/guidelines-for-research-ethics-in-the-social-sciences-humanities-law-and-theology/

Neumann, P. J., Sanders, G. D., Russell, L. B., Ganiats, T. G., & Siegel, J. E. (2016). *Cost-Effectiveness in Health and Medicine*. Oxford University Press.

Newson, A. J., & Dive, L. (2021). Taking seriousness seriously in genomic health. *European Journal of Human Genetics*, 1–2. https://doi.org/10.1038/s41431-021-01002-9

NHS England Cancer Drugs Fund Team. (2016). *Appraisal and Funding of Cancer Drugs from July 2016 (including the new Cancer Drugs Fund)—A new deal for patients, taxpayers and industry*. https://www.england.nhs.uk/wp-content/uploads/2013/04/cdf-sop.pdf

NICE. (2009). *National Institute for Health and Clinical Excellence. Appraising life-extending, end of life treatments.* NICE. https://www.nice.org.uk/guidance/gid-tag387/documents/appraising-life-extending-end-of-life-treatments-paper2

NICE Citizens Council. (2008). *Quality Adjusted Life Years (QALYs) and the Severity of Illness*. National Institute for Health and Care Excellence (NICE). http://www.ncbi.nlm.nih.gov/books/NBK401693/

Nord, E. (1993a). [Health care politicians are not concerned about maximum health gain per crown]. *Tidsskrift for Den Norske Laegeforening: Tidsskrift for Praktisk Medicin, Ny Raekke*, *113*(11), 1371–1373.

Nord, E. (1993b). The trade-off between severity of illness and treatment effect in cost-value analysis of health care. *Health Policy*, *24*(3), 227–238. https://doi.org/10.1016/0168-8510(93)90042-N

Nord, E. (2005). Concerns for the worse off: Fair innings versus severity. *Social Science & Medicine*, *60*(2), 257–263. https://doi.org/10.1016/j.socscimed.2004.05.003

Nord, E., & Johansen, R. (2014). 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*, *116*(2), 281–288. https://doi.org/10.1016/j.healthpol.2014.02.009

Nordenfelt, L. (1995). *On the Nature of Health: An action-theoretic approach* (2nd ed.). Rodopi.

Nordhaug, M., & Nortvedt, P. (2011). Justice and Proximity: Problems for an Ethics of Care. *Health Care Analysis*, *19*(1), 3–14. https://doi.org/10.1007/s10728-010-0159-3

Norheim, O. F. (2010). Priority To The Young Or To Those With Least Lifetime Health? *The American Journal of Bioethics*, *10*(4), 60–61. https://doi.org/10.1080/15265161003697305

Norheim, O. F., Allgott, B., Aschim, B., Førde, R., Gjul, G. K., Gundersen, T., Kakad, M., Kjellevold, A., Kvinsland, S., Melberg, H. O., Moen, A., Mæland, Ø., Olsen, J. A., & Sjøli, S. (2014). *Åpent og rettferdig—Prioriteringer i helsetjenesten* (NOU 2014:12; p. 219). Helse- og Omsorgsdepartementet. https://www.regjeringen.no/no/dokumenter/NOU-2014-12/id2076730/

Norheim, O. F., Baltussen, R., Johri, M., Chisholm, D., Nord, E., Brock, D., Carlsson, P., Cookson, R., Daniels, N., Danis, M., Fleurbaey, M., Johansson, K. A., Kapiriri, L., Littlejohns, P., Mbeeli, T., Rao, K. D., Edejer, T. T.-T., & Wikler, D. (2014). Guidance on priority setting in health care (GPS-Health): The inclusion of equity criteria not captured by cost-effectiveness analysis. *Cost Effectiveness and Resource Allocation*, *12*(1), 18. https://doi.org/10.1186/1478-7547-12-18

Norheim, O. F., Emanuel, E. J., & Millum, J. (Eds.). (2019). *Global Health Priority-Setting: Beyond Cost-Effectiveness*. Oxford University Press. https://doi.org/10.1093/oso/9780190912765.001.0001

Normand, C. (2012). Setting priorities in and for end-of-life care: Challenges in the application of economic evaluation. *Health Economics, Policy and Law*, *7*(4), 431–439. https://doi.org/10.1017/S1744133112000229

Nortvedt, P., Pedersen, R., Grøthe, K. H., Nordhaug, M., Kirkevold, M., Slettebø, A., Brinchmann, B. S., & Andersen, B. (2008). Clinical prioritisations of healthcare for the aged—Professional roles. *Journal of Medical Ethics*, *34*(5), 332–335. https://doi.org/10.1136/jme.2007.020693

Norwegian Ministry & of Health and Care Services. (2017). *Principles for priority setting in health care*. https://www.regjeringen.no/contentassets/439a420e01914a18b21f351143ccc6af/en-gb/pdfs/stm201520160034000engpdfs.pdf

NOU 1987: 23. (1987). *Retningslinjer for prioritering innen norsk helsetjeneste*. Helse- og omsorgsdepartmentet.

NOU 1997: 7. (1997). *Piller, prioritering og politikk Hva slags refusjonsordning trenger pasienter og samfunn?* Helse- og omsorgsdepartementet.

NOU 1997: 18. (1997). *Prioritering på ny: Gjennomgang av retningslinjer for prioriteringer innen norsk helsetjeneste*. Helse- og omsorgsdepartmentet.

NOU 2014: 12. (2014). *Åpent og rettferdig – prioriteringer i helsetjenesten*. Helse- og omsorgsdepartmentet.

NOU 2018: 16. (2018). *Det viktigste først: Prinsipper for prioritering i den kommunale helse- og omsorgstjenesten og for offentlig finansierte tannhelsetjenester*. Helse- og omsorgsdepartementet.

NOU 2023: 4. (2023). *Tid for handling: Personellet i en bærekraftig helse- og omsorgstjeneste*. Helse- og omsorgsdepartementet. https://www.regjeringen.no/contentassets/337fef958f2148bebd326f0749a1213d/no/pdfs/nou202320230004000dddpdfs.pdf

Nussbaum, M., & Sen, A. (Eds.). (1993). *The Quality of Life* (Reprint edition). Clarendon Press.

Oddsson, K. (2003). Assessing attitude towards prioritizing in healthcare in Iceland. *Health Policy (Amsterdam, Netherlands)*, *66*(2), 135–146. https://doi.org/10.1016/s0168-8510(02)00211-7

OECD. (2019). *Health spending set to outpace GDP growth to 2030*. Organisation for Economic Co-operation and Development. https://www.oecd.org/health/health-spending-set-to-outpace-gdp-growth-to-2030.htm#

Olsen, J. A. (2011a). Concepts of Equity and Fairness in Health and Health Care. In S. Glied & P. C. Smith (Eds.), *The Oxford Handbook of Health Economics* (pp. 813–836). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199238828.013.0034

Olsen, J. A. (2011b). Concepts of Equity and Fairness in Health and Health Care. In S. Glied & P. C. Smith (Eds.), *The Oxford Handbook of Health Economics* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199238828.013.0034

Olsen, J. A. (2013). Hva menes med «sykdommens alvorlighetsgrad»? *Tidsskrift for Den norske legeforening*. https://doi.org/10.4045/tidsskr.12.0659

Opsahl, J.-H. (2020). Legers behandling av familie og venner. *Tidsskrift for Den norske legeforening*. https://tidsskriftet.no/2020/09/kronikk/legers-behandling-av-familie-og-venner

Ottersen, T., Førde, R., Kakad, M., Kjellevold, A., Melberg, H. O., Moen, A., Ringard, Å., & Norheim, O. F. (2016). A new proposal for priority setting in Norway: Open and fair. *Health Policy (Amsterdam, Netherlands)*, *120*(3), 246–251. https://doi.org/10.1016/j.healthpol.2016.01.012

Ottersen, T., Mæstad, O., & Norheim, O. (2014). Lifetime QALY prioritarianism in priority setting: Quantification of the inherent trade-off. *Cost Effectiveness and Resource Allocation*, *12*(1), 2. https://doi.org/10.1186/1478-7547-12-2

Paley, J. (2000). Paradigms and presuppositions: The difference between qualitative and quantitative research. *Scholarly Inquiry for Nursing Practice*, *14*(2), 143–155; discussion 157-163.

Parfit, D. (1991). *Equality or Priority*. University of Kansas, Department of Philosophy. https://kuscholarworks.ku.edu/handle/1808/12405

Patton, M. Q. (2015). *Qualitative Research & Evaluation Methods*. SAGE Publications, Inc.

Pellegrino, E. D. (1986). Rationing Health Care: The Ethics of Medical Gatekeeping. *Journal of Contemporary Health Law and Policy*, *2*, 23–46.

Pilgrim, D. (2008). `Recovery' and current mental health policy. *Chronic Illness*, *4*(4), 295–304. https://doi.org/10.1177/1742395308097863

Pinto-Prades, J.-L., Sánchez-Martínez, F.-I., Corbacho, B., & Baker, R. (2014). Valuing QALYs at the end of life. *Social Science & Medicine (1982)*, *113*, 5–14. https://doi.org/10.1016/j.socscimed.2014.04.039

Pitkin, H. (1967). *The Concept of Representation*. University of California Press. https://www-jstor-org.ezproxy.uio.no/stable/2214898

Powers, M. (2005). Bioethics as Politics: The Limits of Moral Expertise. *Kennedy Institute of Ethics Journal*, *15*(3), 305–322. https://doi.org/10.1353/ken.2005.0023

Prah Ruger, J. (2010). Health Capability: Conceptualization and Operationalization. *American Journal of Public Health*, *100*(1), 41–49. https://doi.org/10.2105/AJPH.2008.143651

Prieto, L., & Sacristán, J. A. (2003). Problems and solutions in calculating quality-adjusted life years (QALYs). *Health and Quality of Life Outcomes*, *1*, 80. https://doi.org/10.1186/1477-7525-1-80

Prioriteringscentrum. (2017). *Nationell modell för öppna prioriteringar inom hälsooch sjukvård—Ett verktyg för rangordning* (3rd ed.). Prioteringscentrum. https://liu.diva-portal.org/smash/get/diva2:1144043/FULLTEXT01.pdf

Rawls, J. (1999). *A theory of justice* (Revised). Harvard university press.

Reckers-Droog, V. T., van Exel, N. J. A., & Brouwer, W. B. F. (2018). Looking back and moving forward: On the application of proportional shortfall in healthcare priority setting in the Netherlands. *Health Policy*, *122*(6), 621–629. https://doi.org/10.1016/j.healthpol.2018.04.001

Reich, R. B. (1988). Policy making in a democracy. In R. B. Reich (Ed.), *The Power of Public Ideas* (pp. 123–156). Harvard University Press.

Retriever Norway. (2023). *Atekst – Retriever Norge*. Atekt - Retriever Norge. https://www.retriever.no/tag/atekst/

Reynolds, J. M. (2022). *The Life Worth Living*. University of Minnesota Press.

Richardson, J., Iezzi, A., & Maxwell, A. (2017). How important is severity for the evaluation of health services: New evidence using the relative social willingness to pay instrument. *The European Journal of Health Economics*, *18*(6), 671–683. https://doi.org/10.1007/s10198-016-0817-y

Riksdagsförvaltningen. (1997). *Prioriteringar inom hälso- och sjukvården Socialutskottets Betänkande 1996/97:SoU14—Riksdagen*. https://www.riksdagen.se/sv/dokument-lagar/arende/betankande/prioriteringar-inom-halso--och-sjukvarden_GK01SoU14

Riksdagsförvaltningen. (2018, November 21). *Prioriteringar inom hälso- och sjukvården Kommittédirektiv 1992:8—Riksdagen*. https://www.riksdagen.se/sv/dokument-lagar/dokument/kommittedirektiv/prioriteringar-inom-halso--och-sjukvarden_GGB18

Risdon, A., Eccleston, C., Crombez, G., & McCracken, L. (2003). How can we learn to live with pain? A Q-methodological analysis of the diverse understandings of acceptance of chronic pain. *Social Science & Medicine*, *56*(2), 375–386. https://doi.org/10.1016/S0277-9536(02)00043-6

Robeyns, I., & Byskov, M. F. (2023). The Capability Approach. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Summer 2023). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2023/entries/capability-approach/

Rosenberg-Yunger, Z. R. S., Daar, A. S., Thorsteinsdóttir, H., & Martin, D. K. (2011). Priority setting for orphan drugs: An international comparison. *Health Policy*, *100*(1), 25–34. https://doi.org/10.1016/j.healthpol.2010.09.008

RStudio Team. (2020). *Integrated Development for R* [RStudio]. http://rstudio.com/

Ryan, M., Scott, D. A., Reeves, C., Bate, A., van Teijlingen, E. R., Russell, E. M., Napper, M., & Robb, C. M. (2001). Eliciting public preferences for healthcare: A systematic review of techniques. *Health Technology Assessment (Winchester, England)*, *5*(5), Article 5. https://doi.org/10.3310/hta5050

Ryynänen, O. P., Myllykangas, M., Kinnunen, J., & Takala, J. (1999). Attitudes to health care prioritisation methods and criteria among nurses, doctors, politicians and the general public. *Social Science & Medicine (1982)*, *49*(11), 1529–1539. https://doi.org/10.1016/s0277-9536(99)00222-1

Salazar, M. K. (1990). Interviewer Bias: How it Affects Survey Research. *AAOHN Journal*, *38*(12), 567–572. https://doi.org/10.1177/216507999003801203

Sale, J. E. M., Lohfeld, L. H., & Brazil, K. (2002). Revisiting the Quantitative-Qualitative Debate: Implications for Mixed-Methods Research. *Quality & Quantity*, *36*(1), 43–53. https://doi.org/10.1023/A:1014301607592

Salomon, J. A. (2017). Quality Adjusted Life Years. In S. R. Quah (Ed.), *International Encyclopedia of Public Health (Second Edition)* (pp. 224–228). Academic Press. https://doi.org/10.1016/B978-0-12-803678-5.00368-4

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, *52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Savulescu, J., Gyngell, C., & Kahane, G. (2021). Collective Reflective Equilibrium in Practice (CREP) and controversial novel technologies. *Bioethics*, *35*(7), 652–663. https://doi.org/10.1111/bioe.12869

Scherpenzeel, A. C. (2010). How Representative Are Online Panels? Problems of Coverage and Selection and Possible Solutions. In *Social and Behavioral Research and the Internet*. Routledge.

Scheunemann, L. P., & White, D. B. (2011). The Ethics and Reality of Rationing in Medicine. *Chest*, *140*(6), 1625–1632. https://doi.org/10.1378/chest.11-0622

Schneiderman, L. J. (2011). Rationing Just Medical Care. *The American Journal of Bioethics*, *11*(7), 7–14. https://doi.org/10.1080/15265161.2011.577511

Schoon, R., & Chi, C. (2022). Integrating Citizens Juries and Discrete Choice Experiments: Methodological issues in the measurement of public values in healthcare priority setting. *Social Science & Medicine*, *309*, 115223. https://doi.org/10.1016/j.socscimed.2022.115223

Schurer, M., Matthijsse, S. M., Vossen, C. Y., van Keep, M., Horscroft, J., Chapman, A.-M., & Akehurst, R. L. (2022). Varying Willingness to Pay Based on Severity of Illness: Impact on Health Technology Assessment Outcomes of Inpatient and Outpatient Drug Therapies in The Netherlands. *Value in Health*, *25*(1), 91–103. https://doi.org/10.1016/j.jval.2021.08.003

Schwappach, D. L. B. (2002). Resource allocation, social values and the QALY: A review of the debate and empirical evidence. *Health Expectations*, *5*(3), 210–222. https://doi.org/10.1046/j.1369-6513.2002.00182.x

Sen, A. (1993). Capability and Well-Being. In M. Nussbaum & A. Sen (Eds.), *The Quality of Life*. Oxford University Press.

Shah, K. K. (2009). Severity of illness and priority setting in healthcare: A review of the literature. *Health Policy*, *93*(2/3), 77–84.

Shah, K. K., Tsuchiya, A., & Wailoo, A. J. (2018). Valuing health at the end of life: A review of stated preference studies in the social sciences literature. *Social Science & Medicine (1982)*, *204*, 39–50. https://doi.org/10.1016/j.socscimed.2018.03.010

Shiroiwa, T., Igarashi, A., Fukuda, T., & Ikeda, S. (2013). WTP for a QALY and health states: More money for severer health states? *Cost Effectiveness and Resource Allocation : C/E*, *11*, 22. https://doi.org/10.1186/1478-7547-11-22

Skedgel, C., Henderson, N., Towse, A., Mott, D., & Green, C. (2022). Considering Severity in Health Technology Assessment: Can We Do Better? *Value in Health*, *25*(8), 1399–1403. https://doi.org/10.1016/j.jval.2022.02.004

Solberg, C. T., Barra, M., Sandman, L., & Hoffmann, B. (2023). Severity as a moral qualifier of malady. *BMC Medical Ethics*, *24*(1), Article 1. https://doi.org/10.1186/s12910-023-00903-2

Spencer, A., Rivero-Arias, O., Wong, R., Tsuchiya, A., Bleichrodt, H., Edwards, R. T., Norman, R., Lloyd, A., & Clarke, P. (2022). The QALY at 50: One story many voices. *Social Science & Medicine*, *296*, 114653. https://doi.org/10.1016/j.socscimed.2021.114653

Stainton Rogers, R. (1995). Q methodology. In J. Smith, R. Harre, & L. Van Langenhove (Eds.), *Rethinking methods in psychology*. SAGE Publications, Inc.

Stenmarck, M. S., Engen, C., & Strand, R. (2021). Reframing cancer: Challenging the discourse on cancer and cancer drugs—a Norwegian perspective: Reframing Cancer. *BMC Medical Ethics*, *22*(1), 126. https://doi.org/10.1186/s12910-021-00693-5

Stenmarck, M. S., Jølstad, B., Baker, R., Whitehurst, D. GT., & Barra, M. (2023). A severely fragmented concept: Uncovering citizens' subjective accounts of severity of illness. *Social Science & Medicine*, *330*, 116046. https://doi.org/10.1016/j.socscimed.2023.116046

Stenmarck, M. S., & Nilsen, I. W. (2022). Precision Oncology in the News. In A. Bremer & R. Strand (Eds.), *Precision Oncology and Cancer Biomarkers: Issues at Stake and Matters of Concern* (Vol. 5). Springer International Publishing. https://doi.org/10.1007/978-3-030-92612-0

Stephenson, W. (1935). Technique of Factor Analysis. *Nature*, *136*(3434), Article 3434. https://doi.org/10.1038/136297b0

Stevenson, H. (2016). The Wisdom of the Many in Global Governance: An Epistemic-Democratic Defense of Diversity and Inclusion. *International Studies Quarterly*, *60*(3), 400–412. https://doi.org/10.1093/isq/sqw027

Stewart, E. (2016). *Publics and Their Health Systems: Rethinking Participation*. Springer.

Sullivan, C., & Forrester, M. A. (2018). *Doing Qualitative Research in Psychology: A Practical Guide*. SAGE.

Talbott, A. D. (2010). The Q-Block Method of Indexing QTypologies. *Operant Subjectivity: The International Journal of Q Methodology*, *34*(1), 6–24.

Tenbensel, T. (2010). *Virtual Special Issue Introduction: Public Participation in health policy in high income countries - A review of why, who, what, which and where?* https://doi.org/10.1016/j.socscimed.2010.08.005

Thacher, D. (2015). Perils of Value Neutrality. In M. Kraatz (Ed.), *Institutions and Ideals: Philip Selznick's Legacy for Organizational Studies* (Vol. 44, pp. 317–352). Emerald Group Publishing Limited. https://doi.org/10.1108/S0733-558X20150000044012

The Directorate of Health. (2012). *Økonomisk evaluering av helsetiltak—En veileder*. The Directorate of Health. https://www.helsedirektoratet.no/veiledere/okonomisk-evaluering-av-helsetiltak/%C3%98konomisk%20evaluering%20av%20helsetiltak%20%E2%80%93%20Veileder.pdf/_/attachment/inline/a4c0d8f4-6420-47a5-a052-344a5f2d0e76:637c45a5f29eda29afb9625968759e6fa4c4af61/%C3%98konomisk%20evaluering%20av%20helsetiltak%20%E2%80%93%20Veileder.pdf

The Directorate of Health. (2018). *Prioritering på klinisk nivå*. The Directorate of Health. https://www.helsedirektoratet.no/rapporter/prioritering-pa-klinisk-niva/Rapport%20fra%20prosjektet%20Prioritering%20p%C3%A5%20klinisk%20niv%C3%A5.pdf/_/attachment/inline/00a28251-70ee-459a-bc80-b19268587427:5d327cbcb7bb25cea31ba26613c8fe77a635412c/Rapport%20fra%20prosjektet%20Prioritering%20p%C3%A5%20klinisk%20niv%C3%A5.pdf

The Norwegian Medical Association. (2021). *Etiske regler for leger*. https://www.legeforeningen.no/om-oss/Styrende-dokumenter/legeforeningens-lover-og-andre-organisatoriske-regler/etiske-regler-for-leger/

The Standing Committee on Health and Care Services. (2022, March 30). *Innstilling fra helse- og omsorgskomiteen om Nytte, ressurs og alvorlighet – Prioritering i helse- og omsorgstjenesten* [Inns]. Stortinget; helse- og omsorgskomiteen. https://www.stortinget.no/no/Saker-og-publikasjoner/Publikasjoner/Innstillinger/Stortinget/2021-2022/inns-202122-221s/?all=true

Thomas, Dan B. & Baas, Larry R. (1993). The Issue of Generalization in Q Methodology: 'Reliable Schematics' Revisited. *Operant Subjectivity: The International Journal of Q Methodology*, *16(1/2)*, 18–36.

Toombs, S. K. (1995). The lived experience of disability. *Human Studies*, *18*(1), 9–23. https://doi.org/10.1007/BF01322837

Tsuchiya, A. (2000). QALYs and ageism: Philosophical theories and age weighting. *Health Economics*, *9*(1), 57–68. https://doi.org/10.1002/(sici)1099-1050(200001)9:1<57::aid-hec484>3.0.co;2-n

Twaddle, A. (1994). Disease, illness and sickness revisited. In A. Twaddle & Nordenfelt, L (Eds.), *Disease, Illness and Sickness: Three Central Concepts in the Theory of Health* (pp. 1–18).

Ubel, P. A., Loewenstein, G., Scanlon, D., & Kamlet, M. (1996). Individual utilities are inconsistent with rationing choices: A partial explanation of why Oregon's cost-effectiveness list failed. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, *16*(2), 108–116. https://doi.org/10.1177/0272989X9601600202

Ubel, P. A., Spranca, M. D., Dekay, M. L., Hershey, J. C., & Asch, D. A. (1998). Public preferences for prevention versus cure: What if an ounce of prevention is worth only an ounce of cure? *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, *18*(2), 141–148. https://doi.org/10.1177/0272989X9801800202 van der Weele, S. (2021). Thick Concepts in Social Research: What, Why, and How? *International Journal of Qualitative Methods*, *20*, 16094069211066165. https://doi.org/10.1177/16094069211066165

van Exel, J., Morée, M., Koopmanschap, M., Goedheijt, T. S., & Brouwer, W. (2006). Respite care—An explorative study of demand and use in Dutch informal caregivers. *Health Policy*, *78*(2), 194–208. https://doi.org/10.1016/j.healthpol.2005.11.002

van Exel, N., & de Graaf, G. (2005). *Q methodology—A sneak preview* (p. 30).

van Exel, N. J. A., Koolman, A. H. E., de Graaf, G., & Brouwer, W. B. F. (2011). *Overweight and obesity in Dutch adolescents: Associations with health lifestyle, personality, social context and future consequences*. Erasmus MC.

Venkatapuram, S. (2013). *Health Justice: An Argument from the Capabilities Approach*. John Wiley & Sons.

Walker, R. L., & Siegel, A. W. (2002). Morality and the limits of societal values in health care allocation. *Health Economics*, *11*(3), 265–273. https://doi.org/10.1002/hec.665

Warren, M. E. (2009). Citizen Participation and Democratic Deficits: Considerations from the Perspective of Democratic Theory. In J. DeBardeleben & J. H. Pammett (Eds.), *Activating the Citizen: Dilemmas of Participation in Europe and Canada* (pp. 17–40). Palgrave Macmillan UK. https://doi.org/10.1057/9780230240902_2

Watts, S., & Stenner, P. (2022, April 10). *Doing Q Methodological Research*. SAGE Publications Ltd. https://uk-sagepub-com.ezproxy.uio.no/en-gb/eur/book/doing-q-methodological-research

Weinstein, M. C., Torrance, G., & McGuire, A. (2009). QALYs: The Basics. *Value in Health*, *12*, S5–S9. https://doi.org/10.1111/j.1524-4733.2009.00515.x

Wester, G., & Bringedal, B. (2018). The Norwegian National Council for Priority Setting in Health Care: Decisions and justifications. *Health Economics, Policy, and Law*, *13*(2), 118–136. https://doi.org/10.1017/S1744133117000020

Whitehead, S. J., & Ali, S. (2010). Health outcomes in economic evaluation: The QALY and utilities. *British Medical Bulletin*, *96*(1), 5–21. https://doi.org/10.1093/bmb/ldq033

Whitty, J. A., Lancsar, E., Rixon, K., Golenko, X., & Ratcliffe, J. (2014). A Systematic Review of Stated Preference Studies Reporting Public Preferences for Healthcare Priority Setting. *The Patient - Patient-Centered Outcomes Research*, *7*(4), 365–386. https://doi.org/10.1007/s40271-014-0063-2

Wichmann, A. B., Goltstein, L. C. M. J., Obihara, N. J., Berendsen, M. R., Van Houdenhoven, M., Morrison, R. S., Johnston, B. M., & Engels, Y. (2020). QALY-time: Experts' view on the use of the quality-adjusted life year in cost-effectiveness analysis in palliative care. *BMC Health Services Research*, *20*(1), Article 1. https://doi.org/10.1186/s12913-020-05521-x

Williams, A. (1988). Priority setting in public and private health care. A guide through the ideological jungle. *Journal of Health Economics*, *7*(2), 173–183. https://doi.org/10.1016/0167-6296(88)90015-x

Williams, A. (1997a). Intergenerational Equity: An Exploration of the 'Fair Innings' Argument. *Health Economics*, *6*(2), 117–132. https://doi.org/10.1002/(SICI)1099-1050(199703)6:2<117::AID-HEC256>3.0.CO;2-B

Williams, A. (1997b). Intergenerational Equity: An Exploration of the 'Fair Innings' Argument. *Health Economics*, *6*(2), 117–132. https://doi.org/10.1002/(SICI)1099-1050(199703)6:2<117::AID-HEC256>3.0.CO;2-B

Williams, I., Dickinson, H., & Robinson, S. (2012a). *Rationing in Health Care: The Theory and Practice of Priority Setting*. Policy Press.

Williams, I., Dickinson, H., & Robinson, S. (2012b). *Rationing in health care: The theory and practice of priority setting*. Policy Press.

Williams, I., Phillips, D., Nicholson, C., & Shearer, H. (2014). Evaluation of a deliberative approach to citizen involvement in health care priority setting. *Leadership in Health Services*, *27*(1), 5–19. https://doi.org/10.1108/LHS-01-2013-0002

Williamson, L. (2014). Patient and citizen participation in health: The need for improved ethical support. *The American Journal of Bioethics: AJOB*, *14*(6), 4–16. https://doi.org/10.1080/15265161.2014.900139

Winkelhage, J., & Diederich, A. (2012). The Relevance of Personal Characteristics in Allocating Health Care Resources—Controversial Preferences of Laypersons with Different Educational Backgrounds. *International Journal of Environmental Research and Public Health*, *9*(1), Article 1. https://doi.org/10.3390/ijerph9010223

Wisløff, T. (2015). Priority-setting criteria in the Norwegian health services. *Tidsskrift for Den Norske Legeforening*. https://tidsskriftet.no/en/2015/08/priority-setting-criteria-norwegian-health-services

World Health Organization. (2010). *The world health report: Health systems financing: the path to universal coverage*. World Health Organization. https://apps.who.int/iris/handle/10665/44371

World Health Organization. (2022). *The ICD-11 classification of mental and behavioural disorders.* https://icd-who-int.ezproxy.uio.no/en

Wyller, V. B. (2014). Give to the doctor what is due to the doctor! Why «fair rationing at the bedside» is impossible. In M. Danis, S. A. Hurst, L. Fleck, R. Forde, & A. Slowther (Eds.), *Fair Resource Allocation and Rationing at the Bedside*. Oxford University Press.

Wyller, V. B. (2017). The bedside rationing paradigm and the shortcomings of modernist ethics. *Clinical Ethics*, *12*(2), 70–75. https://doi.org/10.1177/1477750917691888

Yew, V. W. C., & Noor, N. A. M. (2014). *ANTHROPOLOGICAL INQUIRY OF DISEASE, ILLNESS AND SICKNESS*. *9*(2).

Zimmerman, M., Morgan, T. A., & Stanton, K. (2018). The severity of psychiatric disorders. *World Psychiatry*, *17*(3), 258–275. https://doi.org/10.1002/wps.20569

Appendices

Appendix A: Group interview guide

Appendix B: Questionnaire (for group interviews and Q sort exercises)

Appendix C: Condition of Instruction (for Q sort exercises)

Appendix D: Post-sort questions (for Q sort exercises)

Appendix A: Group interview guide

QWS Interview guide – WP1

Subjective views on severity

Welcome and practical information

Welcome

Health and safety: Where are the toilets? Escape routes. Option to leave the conversation.

Practical: 3 hours planned. The plan for food/drinks. Distribution of papers and pens. Remuneration.

Ethics/Consent

- Distribution and review of consent form.
- Mobile on silent/flight mode, no use of audio recordings or pictures.
- Confidentiality.

Read: Thank you for being willing to contribute to this project. By participating in this focus group you will contribute to research on health priority setting in Norway. During the conversation some might experience that topics arise that are uncomfortable to discuss or that bring feelings of anxiousness. We understand that this can be difficult and want to emphasise that you can leave the group if you wish to. *The conversation will be recorded and then transcribed, and anonymised*. These data will be securely stored and the anonymised material will only be available for the project collaborators, *as described in the consent form*. *[CHECK CONSENT FROM EVERY PARTICIPANT]*. You have a duty of confidentiality concerning what you learn about other participants, and it is prohibited to record sound or images of each other during the conversation.

[Let participants ask questions.]

Introduction to the group conversation on *SEVERITY*

Introduce the theme

Norwegian priority setting guidelines state that the cost of treatment and the benefit, i.e. the health benefit, for the patient are important factors that should be taken into account. In addition we have a *severity criterion*.

In simplified terms, the severity criterion states that we should prioritise saying ‘yes’ to treatments that are really “too expensive” if the condition that is treated is “severe enough”. *[pause]*

This also means that one should say ‘no’ to treatment for less severe conditions if it is too costly. *[pause]*

We don’t know exactly what the population think *severity* is. We also don’t know what role the population thinks severity should have in health priority setting.

Guidelines for the conversation

Before you begin to discuss severity, we would like to emphasise that this is a topic with a lot of different viewpoints. We do not think there are any right or wrong views on this. Health personnel, economists, and philosophers also discuss what severity is, and what role severity should play in priority setting.

The purpose of this focus group is to explore the different *subjective* viewpoints in the population in Norway. All thoughts and input, thought through or spontaneous, clear ideas or vague impressions, based on personal or professional experience, are equally useful, important, and welcome today.

We want you to talk to each other rather than to us and we will first and foremost be moderators. We will nonetheless participate in the conversation, and can come with input and interpretations, or follow-up questions if we wish to hear more about something.

So…*what do* you *think severity is? Begin with the first associations that come to mind, and then we can talk more about priority-setting later on*…

Appendix B: Questionnaire (for group interviews and Q sort exercises)

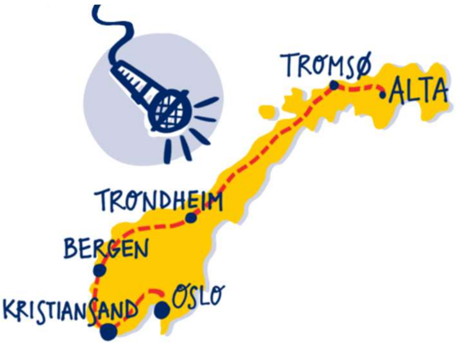
Conversation on Severity

Thank you for filling out this form. All the information we collect is for research purposes. Information will be unidentified and deleted after the project ends. The purpose of this form it to be able to describe the participants as a group, and to better be able to interpret the meaning of the conversation you have participated in.

Questionnaire number (RID):

Date:

Project collaborator:



I CONSENT TO RESPOND TO THIS QUESTIONNAIRE AND TO THAT MY RESPONSES ARE STORED AND USED AS DESCRIBED.
(THIS CONSENT IS NOT VALID WITHOUT A SIGNED CONSENT FORM FOR THE GROUP CONVERSATIONS ABOUT SEVERITY.)

Place and date

Participant signature

Participant name in capital letters

QUESTIONNAIRE

PART I – ABOUT YOU

Questions 1–12 allow us to describe the participants in these group conversations. In addition, this information helps us to recruit future participants. In this phase of the study, we wish to meet as many different people as possible, and in this way we identify the breadth in the participants' background. These responses also help us to interpret the meanings within the conversations. Your responses will be saved unidentified and separate from directly identifiable data like name and contact information.

S1 Sex

Female	Male	Other/Prefer not to say
☐	☐	☐

S2 Age

☐ 18-30
☐ 31-50
☐ 51-66
☐ 67+
☐ Prefer not to say

S3 What is your highest education level
(One answer only.)

☐ Elementary school
☐ Upper secondary school
☐ Fagbrev
☐ Bachelor degree
☐ Master degree
☐ Profession study (MD, dentist etc.)
☐ PhD
☐ Other (please describe)
☐ Prefer not to say

S4 If you live with someone, state the number within each age category.

Children (0 – 16 years old)	
Adults (17 – 66 years old)	
Elderly (> 67 years old)	

☐ I am married/cohabitant

☐ Prefer not to say

S5 Do you consider yourself religious or spiritual?

(Mark the alternative that best suits you.)

Religious/spiritual *Active in a community*	Religious/spiritual *Not active in a community*	No	Prefer not to say
☐	☐	☐	☐

S6 Have you or anyone you know had a severe illness?

(Several crosses are possible.)

☐ Yes, passing

☐ Yes, chronic

☐ Yes, with deadly outcome

☐ Don’t know

☐ No

☐ Prefer not to say

If you responded ‘Yes’ to **Question 6**, please elaborate in **Question 6b**

S6b Which relation did you have to that person?

(Several crosses are possible.)

☐ Partner

☐ Parent

☐ Child

☐ Sibling

☐ Friend

☐ Other family member (please describe)

☐ Other (please describe)

☐ Prefer not to say

S7 **How do you consider your own health?**
(One answer only.)

Very good	Good	Okay	Poor	Very poor	Don’t know	Prefer not to say
☐	☐	☐	☐	☐	☐	☐

S8 **Generally, what is your impression of the quality of the Norwegian public healthcare system?**
(One answer only.)

Very good	Good	Okay	Poor	Very poor	Don’t know	Prefer not to say
☐	☐	☐	☐	☐	☐	☐

PART II – ABOUT THE CONVERSATION

S9 **What would you say has had the greatest influence on your view of what *severity* means in relation to illness?**
(Use the numbers 1, 2 and 3 to rank the most important ones.)

☐ My upbringing
☐ Media
☐ Religious views/Outlook
☐ Family
☐ Education
☐ Political views
☐ Life experience
☐ Prefer not to say

☐ Work/profession
☐ Single event in life (please describe)
☐ Organisations (please describe)
☐ Other (please describe)
☐ Don’t know/Diffucult to say

We will not identify you when we report on our findings in this study, but we wish to convey something about the background of participants to better illustrate the subjective viewpoints that have been expressed in these meetings.

S10 **If you have any suggestions to how we can describe you, you can write it in the section below. You can tie the description to what you might think has had a large influence on what you have expressed in the group conversation. What would you say is your field of expertise? Is there something else that it is important to know?**
(For example: 'disabled nurse', 'lawyer and mother', 'man in his thirties med responsibility of sick father'.)

S11 **We would like to speak to as many people as possible. The goal is to know as many viewpoints on severity as possible. Er det 'someone' you think might have valuable perspectives on this?**
(For example: 'minimum pensioners', 'immigrant', 'cancer patients')

S12 **How do you think the group conversation went?**
(Did the facilitator participate too much or too little? Were examples that the facilitators mentioned useful? Did you feel that you got the chance to speak?)

Thank you for your help!

Appendix C: Condition of Instruction (for Q sort exercises)

Q-Sorting Condition of Instruction (CoI) – WP1

Subjective views on severity

Welcome and practical information

Welcome

Introduction and 'housekeeping': Facilities, escape routes. Option to leave the conversation.

Practical: 3 hours planned. The plan for food/drinks. Distribution of papers and pens. Remuneration.

Ethics: Phone on silent/flight mode, prohibited to record audio or video. Confidentiality.

Consent

Check consent. Let participants ask questions.

Ethics

Thank you for being willing to contribute to this project. By participating in this focus group you will contribute to research on health priority setting in Norway. During the exercise some might experience that topics arise that are uncomfortable to discuss or that bring feelings of anxiety. We understand that this can be difficult and want to emphasise that you can leave the group if you wish to. The conversation after the exercise will be recorded and then transcribed, and anonymised. These data will be securely stored and the anonymised material will only be available for the project collaborators on SEVPRI.

Sorting exercise on SEVERITY

Introduce the topic

This is a research project that is about how we should use resources in the healthcare system in the best and fairest way. Specifically, we explore the term 'severity': what does it means that something is severe in the context and health and illness. Amongst professionals, there is agreement that how severe an illness is should mean something for how the illness is prioritised. But we don't know exactly what the population think *severity* is.

We would like to emphasise that this is a topic where there can be a lot of different viewpoints. We do not think there are any right or wrong views on this. What severity is, is also a topic that there is no “professional” consensus on. Health personnel, economists, and philosophers still discuss what severity is, and what role it should play in priority setting.

We have travelled around Norway and spoken to members of the general population about what they think severity is. From these conversations we have collected a set of statements from the participants on what severity means to them. We now want to know what you think about severity.

Are there any questions so far?

Condition of Instruction (CoI)

On the cards you have in front of you there are things people have said about severity in the context of health and illness. You might agree with some of the cards. Other you will disagree with. There are no right or wrong answers here. We ask you to read the cards carefully and put them in three piles: agree, disagree, and statements you feel neutral about or don’t have an opinion on. You don’t need to divide them perfectly now, I will ask you to sort them in more detail later and you can make changes as we go. When you have finished, you can write down the number of cards in each pile (agree, disagree, neutral) on the paper next to you.

For every statement we have one question: do you agree or disagree with the statement?

Give instructions that this part of the exercise should be completely quite quickly. Sort on gut instinct.

Agree statements

Now, take the pile with the statements form “agree” and find the 1 statement you agree the *most* with. Place this in the column furthest to the right.

Go back to the pile and choose the 1 statement you agree with the most. Place this in column number 2 from the right.

Go back to the pile and choose the 3 statements you agree with the most. Place these statements in column number 3 form the right.

Disagree statements

Now we will do the same task with statements you disagree with. So, take the pile with the statements form “disagree” and find the 1 statement you disagree the *most* with. Place this in the column furthest to the left.

Go back to the pile and choose the 1 statement you disagree with the most. Place this in column number 2 from the left.

Go back to the pile and choose the 3 statements you disagree with the most. Place these statements in column number 3 form the left.

Keep placing the rest of the cards from the “disagree” pile like before, in the order you disagree with them, and work your way towards the middle of the grid until you have no more cards left.

The middle

Go back to the “agree” pile. Complete the placement of the “agree” cards like before, in the order you agree with them. Work your way from the right towards the middle of the grid, until you have no more cards left.

Finally, go back to the pile with the cards you feel more neutral about, or don’t have a strong opinion on. Place these cards where there are open spaces in the grid, depending on how much you agree or disagree with them. So, begin with those you disagree with the most and place them towards the left, and finish with the ones you agree with the most towards the right.

After all the cards are placed, ask the participants: are you pleased with the cards the way they are placed now, or would you like to make any changes?

Appendix D: Post-sort questions (for Q sort exercises)

Post-sort questions

1. If you were to sum up your view on ‘severity’ in the context of health and illness, what would you say?

2. Look at the statements to the very right in your diagram (+6). Why did you place these here? What thoughts do you have when you see those statements?

3. Look at the statements to the very left in your diagram (-6). Why did you place these here? What thoughts do you have when you see those statements?

4. The area in the diagram where the statements are presumed to be neither particularly similar or dissimilar to your viewpoint is the ‘0’-column. Can you say which column you feel contains the statements that are neither similar or dissimilar to your viewpoint, if it’s a different column (further to the right or left)?

5. Were there any viewpoints that you felt were missing in the statements? Why are those viewpoints important to you?

I

"*It's hard to say anything definitive about what severity really is*": Lay conceptualisations of severity in a healthcare context

Mille Sofie Stenmarck[1, 2*]
David GT Whitehurst[3]
Hilde Lurås[1, 2]
Jorun Rugkåsa[1,4,5]

Affiliations

[1] The Health Services Research Unit – HØKH, Akershus University Hospital HF, Lørenskog, Norway
[2] Faculty of Medicine, University of Oslo, Oslo, Norway
[3] Faculty of Health Sciences, Simon Fraser University, Vancouver, Canada
[4] Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway
[5] Faculty of Health and Social Sciences, University of South-Eastern Norway, Porsgrunn, Norway
[*] Corresponding author (email address: msstenmarck@gmail.com)

Abstract

Background: Demand for healthcare outweighs available resources, making priority setting a critical issue. 'Severity' is a priority-setting criterion in many healthcare systems, including in Norway, Sweden, the Netherlands, and the United Kingdom. However, there is a lack of consensus on what severity means in a healthcare context, both in the academic literature and in policy. Further, while public preference elicitation studies demonstrate support for severity as a relevant concern in priority setting, there is a paucity of research on what severity is taken to mean for the public. The purpose of this article is to explore how severity is conceptualised by members of the general public.

Methods: Semi-structured group interviews were conducted from February to July 2021 with members of the Norwegian adult public (*n*=59). These were transcribed verbatim and subjected to thematic analysis, incorporating inductive and deductive elements.

Results: Through the analysis we arrived at three interrelated main themes. A subjective theme centered on severity as inherently subjective and personal. Emphasis was on the individual's unique insight into their illness, and there was a concern for determining severity fairly for the individual. In a second theme, severity was seen as objective and extrapersonal. Severity was determined by objective criteria, and if deemed severe a condition was equally severe for any person suffering from it. Here, there was a concern for determining severity fairly within and across patient groups. In a final situational theme, severity centered on the second-order effects of illness. These included effects on the individual, such as their ability to work and enjoy their hobbies, effects on those surrounding the patient, such as next of kin, and effects at a societal level, such as production loss. We also identified a concern for determining severity fairly at a societal level.

Conclusions: When operationalising severity in priority-setting frameworks, policymakers appear to have taken for granted that severity is a concrete, well-defined notion. Our findings suggest severity is polyvalent, involving additional concerns for the citizenry beyond those represented in policy. There thereby appears to be considerable dissonance between the views of the public and policy operationalisations of severity.

Keywords: priority setting, health policy, severity, thematic analysis, views of the public.

Background

The demand for healthcare services outweighs available resources, and governments face complex dilemmas of healthcare prioritisation (1,2). Priority setting in healthcare is an issue in both low- and high-income countries, and in publicly-funded and private healthcare systems (3). Healthcare systems rely on priority-setting frameworks to guide decision-making, and there is a broad field of research on the principles underpinning these frameworks. In most universal healthcare systems, priority-setting principles are typically centered on cost-effectiveness criteria (1). 'Severity' is another criterion and has been adopted in several countries, including Norway (4), Sweden (5), the Netherlands (6), and the United Kingdom (UK) (7). Despite the widespread use of severity as a criterion there is a lack of consensus on how to operationalise it.

A severity criterion modifies decision rules in cost-effectiveness analyses, potentially allowing for the recommendation of treatments (for conditions considered to be 'severe') that would otherwise not have met cost-effectiveness thresholds. As such, severity has been described as an ethical decision-modifier (8). In Norway, priority-setting decisions should be based on the three criteria of health benefit, resources, and severity (9). The three criteria are intended to be applied throughout the healthcare system, from health policy to the clinical level, and weighed against each other. The severity criterion is operationalised as absolute quality-adjusted life year (QALY) shortfall (4). The QALY is a health metric that combines quality and quantity of life in a single outcome (10), and 'absolute QALY shortfall' represents the expected loss of QALYs due to illness (11). Other jurisdictions use different operationalisations: the Netherlands operationalise severity as proportional QALY shortfall, calculated by the expected loss of QALYs relative to remaining life expectancy (6); in the UK, a combination of absolute and proportional QALY shortfall is implemented (7); and Sweden employs a severity framework which measures severity according to a qualitative ranking of severity levels, from 'low' to 'very high' (12,13). While severity is a common priority-setting criterion, it is evident that there is no consensus on how to operationalise it in policy.

While attempts at defining severity in the academic literature are usually based on QALYs (14,15), different conceptualisations exist (16–18). Olsen argues that (at least) four different approaches can be identified: severity understood according to (i) how poor one's health is; (ii) short remaining lifetime; (iii) poor prognosis; or (iv) the size of the health loss (11). This academic ambiguity is present also among stakeholders within the healthcare system. Magnussen and colleagues distributed a survey among healthcare personnel, leaders at different levels of the healthcare system, and patient organisations, and find that there is no agreement on what severity means (4). This lack of a shared understanding further complicates the use of the term in a policy context (11).

Decisions on healthcare prioritisation inevitably involves allocating resources to some groups over others, making priority setting in healthcare contentious (1). Because the outcome of these decisions is consequential to the public, knowledge of public views is critical (19). In a literature review of public preference elicitation studies on the relevance of severity in healthcare, Shah demonstrates that severity is considered an important and relevant concern for priority setting across multiple populations (18). Studies conducted across Norway, Denmark, Finland, the UK, and the US establish support among the general public for severity (20–24). Both an Australian, an Icelandic, and a UK study find that general public respondents prefer at least equal priority to the severely ill (25–27). And in a Canadian survey with general public participants, severity was ranked as the most important concern across all respondent groups (28).

While previous literature has sought to elicit the degree to which the public consider severity to be relevant, via methods comprising small-sample focus groups through to population surveys (18), they do not explore what severity is taken to mean. Furthermore, there is a lack of consensus on how to define severity (the aforementioned studies apply different definitions, if they provide one at all), whether for the reader or the participants in the respective studies. Broqvist and colleagues adopt a more explorative approach to understanding severity by comparing views of the public on severity levels within the Swedish priority-setting system (12). They find that the citizenry considers a multitude of different aspects relevant to determining severity, such as physical or psychological impairment, risk of death, and duration of illness. Their findings suggests that severity is interpreted as something more than QALY shortfall—but do not tell us what severity itself is taken to mean, or how these views compare to current policy operationalisations of severity.

It is evident that there is ambiguity surrounding severity in health policy, in the priority-setting literature, and in multiple public preference elicitation studies (16,29). There is also a paucity of research on what severity is taken to mean by members of the public. To address this latter knowledge gap, we conducted group interviews with members of the Norwegian population.

Methods

Design

This study was part of the SEVerity and PRIority setting in healthcare (SEVPRI) project, which seeks to explore the general public's views on severity. Other phases of the SEVPRI project comprise a Q-methodology study to examine accounts of severity and locate shared viewpoints (30), and a cross-sectional survey study to explore the prevalence of different views in a representative sample of the Norwegian population (forthcoming). In the present study, we present a qualitative thematic analysis of group interview data conducted with members of the public to gain a more nuanced and in-depth understanding of severity. Data was collected through group interviews across Norway, in the format of open conversations.

We anticipated that approximately 60 participants were needed to reach saturation, understood as data redundancy (31) in that new data becomes repetitive of what has been expressed in previous conversations (32). To ensure a diversity of perspectives, we sampled purposively to achieve representation of different demographics, including age, education level, socioeconomic background, health status, and region. Ahead of data collection, an introductory text was prepared to introduce participants to the topic of severity (see Additional file 1). This explained, in lay terms, the three priority-setting criteria used in the Norwegian healthcare system, and that particularly severe conditions can be prioritised. The text also stated that our purpose was to explore participants' subjective viewpoints, and that all thoughts, perceptions, and input was welcome. The presentation was kept brief to avoid influencing participants' views or priming them on certain perspectives. It also emphasised that there are no right or wrong answers on this topic, and that health personnel, health economists, and philosophers have disagreed on what severity means.

We aimed to identify the breadth of views on severity, and conversations were therefore not moderated to reach consensus. Rather, the facilitators sought to capture the various views participants held by starting the conversations with an open question, and allowing participants to discuss freely and as uninterrupted as possible. This open conversation style was supported by a topic guide to ensure the same topics were covered across the conversations. The topic guide contained possible attributes of severity considered relevant, compiled following a comprehensive literature search on the subject (16). This included issues like age, death, and pain, according to which we would ask participants if they, for example, felt that risk of death or someone's age made a condition more severe. The conversation format, introductory text, and topic guide were piloted with a user panel at Akershus University Hospital, consisting of eight people from different demographic backgrounds. Following positive feedback and minor linguistic edits, data collection commenced.

Data collection

Participants were recruited via SEVPRI's social media accounts (Facebook and Twitter), sharing a link to an online recruitment platform. Recruitment posters were hung in shops and on lampposts in Oslo, as well as in the waiting rooms of general practitioners in Oslo and Bergen. The recruitment period was February to July 2021. We began sampling widely across the population. To achieve the desired representation of participant characteristics, sampling become increasingly more targeted (e.g., seeking out male participants when the sample became over-represented by women). Due to SARS-CoV-2 pandemic restrictions, the first 14 conversations (February to March 2021) were conducted online, using Zoom (33). The final seven conversations were conducted in-person following the lifting of restrictions (May to July 2021). In-person conversations were conducted across five locations (Oslo, Bergen, Trondheim, Tromsø, and Alta). Details of conversation format and participants are provided in Table 1. Group size was determined based on two considerations: (i) enough participants to have a meaningful discussion, and (ii) not too many participants to allow everyone to voice their views, with time to explore the depths of these views. The nature of online meetings—with time lags, less non-verbal communication, and other digital challenges—made these challenging to moderate with larger groups. Online conversations were therefore conducted with a minimum of two participants and a maximum of four. The in-person conversations had an upper limit of eight participants. Participants received a universal gift card as compensation for participation, with NOK 250 (~€23) for participation in online conversations and NOK 500 (~€45) for in-person conversations, as these required travel to the meeting locale.

Table 1 Overview of conversation format and number of participants (not including facilitator).

Conversation	**Format**	**Number of participants**
C1	Online	2
C2	Online	2
C3	Online	2
C4	Online	2
C5	Online	2
C6	Online	2
C7	Online	1[a]
C8	Online	3
C9	Online	2
C10	Online	2
C11	Online	2
C12	Online	3
C13	Online	3
C14	Online	2
C15	In-person	2
C16	In-person	5
C17	In-person	4
C18	In-person	4
C19	In-person	5
C20	In-person	3
C21	In-person	6

[a] In conversation #7 (C7), a second scheduled participant did not attend. *Anonymised's* Principal Investigator participated in the conversation with the participant; the Principal Investigator's comments were not included in the analysis.

We considered that data saturation had been reached after 21 conversations, with a total of 59 participants. Participant demographics are presented in Table 2. The conversations lasted approximately two hours when online and three hours when in-person. The lead author (MSS) facilitated ten of the conversations and was present in another five. As data collection for SEVPRI was a joint effort by the research team, the remaining six conversations were conducted by two other, non-author members of the SEVPRI research team (Mathias Barra and Odd Borgar Jølstad).

Table 2 Participant demographics from questionnaires. Values are numbers (percentages).

Characteristic	**Participants (n = 59)**
Age category (years)[a]	
18-30	9 (15)
31-50	13 (22)
51-66	24 (41)
67+	11 (19)
No response	2 (3)
Gender	
Female	38 (64)
Male	19 (32)
Other/prefer not to say	2 (3)
Do you consider yourself religious or spiritual?	
Religious and/or spiritual: active in a congregation	11 (19)
Religious and/or spiritual: not active in a congregation	14 (24)
Neither religious nor spiritual	33 (56)
No response	1 (2)
What is your highest completed education level?	
Elementary/Upper secondary (up to 19 years of age)	9 (15)
Undergraduate degree/Apprenticeship	21 (36)
Graduate degree/PhD	27 (46)
No response	2 (2)
Have you or anyone you know well had severe illness?[b]	
Transient	18 (31)
Chronic	30 (51)
Deadly outcome	42 (71)
No response	0 (0)
How do you view your own health?	
Very good/Good	37 (63)
Just fine	15 (25)
Bad/Very bad	6 (10)
No response	1 (2)

---

[a] Age was given in one of the listed age brackets.

[b] Categories are not mutually exclusive.

Following the introductory text, conversations were initiated by an open question, asking: *"what does severity mean to you? Feel free to start with the first associations you have"* [English translation]. Each participant was given the opportunity to respond, followed by a group conversation to explore the various views that emerged as well as items from the topic guide. The facilitators encouraged participants to speak freely, focusing on views participants spontaneously brought up and the conversations that developed organically. Participants were probed to expound on their views for clarification. For example, a facilitator could say: *"you said that age seems relevant to you, what do you mean by that?"* or *"is the condition more severe if it affects a young or an old person?"*. They were at times also provided with examples of how their views would translate into real-world situations.

Following the conversation, participants in the in-person groups completed a questionnaire that asked questions about socioeconomic status, health status, and situations that may have affected their views on severity (see Additional file 2). Online participants completed the same questionnaire over the phone with the facilitator. All conversations were audio recorded and transcribed verbatim by members of the research team, in Norwegian.

Analysis

The data were subjected to qualitative thematic analysis of repeated cycles of induction and deduction (34). Data were stored and coded using NVivo (release 1.6.1). The analysis was conducted in four stages combining different analytic techniques, as outlined in Table 3. Each stage was led by the lead author (MSS), with contributions from and discussions with all co-authors. In step 2, three authors (MSS, JR, HL) separately read three transcripts and identified potential codes, which were then compared and developed into a coding framework.

The aim of the analysis was to elicit the breadth of views expressed in the conversations to identify broad themes across participants, rather than individual views. As such, the themes do not represent specific groups of participants and one participant's contribution might fall into more than one theme. We use quotes from the conversations to illustrate the themes, identified by an alias and a code (C1 through C21) indicating in which group conversation the quote was collected from.

Table 3 The four stages of analysis.

Stages	**Description of the analytical process**
1: From chaos to codes: Read-through	Getting familiar with the data by reading through all transcripts. Note-taking using mind-maps to record topics for potential codes.
2: Coding the material: Deductive-inductive cycles	Three authors (*Anonymised*, *Anonymised*, *Anonymised*) independently coding three manuscripts to ensure quality and congruence of coding. Subsequently coding all transcripts, adapting the codebook as necessary. Dynamically developing codebook during the coding process (inductive).
3: From code to meaning: Identifying themes	Studying the codes in isolation and in conjunction with each other, searching for themes. Creating mind maps of potential themes and identifying if and how codes fit within these.
4: From de-contextualisation to recontextualisation: Descriptions	Connecting the themes to broader body of literature, looking for connections within and between themes. Recontextualising by returning to transcripts to consider if themes reflect what participants discussed. Writing out narrative within themes.

In Stage 3 we arrived at three themes, which we see as representing three interrelated conceptualisations of severity: (1) severity as subjective experience; (2) severity as objective fact; and (3) severity as situation dependent.

Ethics

The Regional Committee for Medical and Health Research Ethics deemed the SEVPRI study outside the remit of the Norwegian Health Research Act (ref. no. 186284). Ethical approval was granted by the Data Privacy Officer at Akershus University Hospital Trust following a detailed Data Protection Impact Assessment (*PVO. Nos 20_200* and *21_200*). All names of people and organisations have been deleted or altered. Participants gave written and oral consent prior to the in-person conversations, and oral consent prior to the digital ones. They were informed that they could withdraw at any stage, but no participants availed of this.

Results

As a backdrop to the results, it was evident that the participants had a lot to say about severity and appeared eager to share their views, and the conversations yielded nuanced and differing perspectives. Following the opening question on what severity meant to them, participants spontaneously associated the term with a multitude of issues, including how they thought age, death, pain, dignity, desert, stigma, next of kin, hope, fear, acuteness, and adaptability were relevant to assess the severity of conditions. Participants seemed reluctant to describe any condition or situation as 'non-severe'. When prompted to specify such circumstances, a few participants volunteered examples such as passing knee injury or cosmetic surgery. In general, there seemed to be a reluctancy throughout the conversations to specify any conditions as definitively without potential of being severe.

Severity as subjective experience

A common topic in the conversations focused on how severity related to the individual's experience of their situation and illness. As such, severity was expressed as an inherently subjective and personal notion, and no condition could be considered severe (or not) until experienced by the individual as such. This precluded a universal judgment of severity, and severity was portrayed as something intrapersonal which should be decided by the individual.

> James (C21): 'I think of severity as a very subjective description of how you experience a condition.'

> Lisa (C19): 'Severity is an individual question and it's an individual assessment.'

It appeared that severity could not be implied from a diagnosis or characteristics of a condition, but from how the patient experiences it. The severity of a condition, such as asthma or a broken leg, might vary between individuals suffering from it, depending on how they perceive their situation. When responding to the opening question, Mary (C18) pointed to this notion of severity as relative:

> ‘All illness is very subjective. What feels severe? For some it’ll be catastrophic to break a leg and immediately feels very severe, if that person thinks that right now my life is ruined because I broke my leg. While for some it’s severe [only] when you’re on your deathbed.’

Steven (C2) further argued that severity related to how the individual perceives their situation. Therefore, neither policymakers nor healthcare professionals could understand the severity of a condition the way a patient does:

> ‘To me that [illness that effects quality of life] would be a severe disease. Even if it wouldn’t have been defined as a severe disease from the authorities or from the healthcare system it would…for me it would be a severe disease because it keeps me from doing, or being part of, of things, so then it’s severe for me.’

Anna (C15) suggested the same by referring to schizophrenia and argued that outsiders cannot fully understand the implications of such a disorder. The power of defining its severity should therefore not rest with doctors, academics, or policymakers, but with the patient:

> ‘You probably have little understanding of the severity [of schizophrenia] if you haven’t felt it in your own body.’

Some expressed skepticism towards a standardised, “one-size-fits-all” approach. Susan (C19) suggested that, following the different interpretations of severity that had been discussed in the conversation, guidelines and standards could not account for the complexity of severity:

> ‘I don’t see how one could set standardised routines to evaluate severity […] With everything we’ve touched on today, so many factors playing into what severity is, I don’t see how one could make a framework that would fit the best for the majority [of situations]. I’m sure there are some sharper minds than mine who can imagine one, but illness and health and severity is as…I mean, there are as many expressions of that as there are people and conditions combined.’

Given the emphasis on the individual’s right to decide what is severe for them, external determination of severity was perceived to impose a form of injustice on the individual. Fair decisions about severity should therefore be done in a manner that is fair to the individual:

> Jennifer (C18): ‘For me, I think I want ownership of my severity. [Severe illness] isn’t something where someone else can say it’s not dangerous, it’ll pass. I think there’s too much of that. It’s about respecting the other’s severity […] It’s about taking the other’s severity seriously. We can’t define it away.’

> Melissa (C20): ‘Depriving people of the subjective experience of severity…you can’t take that away from people. [The subjective severity] is always there. And that’s what the healthcare system has to deal with. The severity that the individual experiences in their situation.’

Severity as objective fact

In other parts of the conversations, severity was conceptualised from an extrapersonal position, independently of individual experience. The severity of, for example, a stroke appeared to depend on elements of that condition, such as risk of death or prognosis. If deemed severe, a stroke would be equally severe for anyone suffering one. Eric (C13) explained that he has a tendency to overestimate the severity of his ailments, and suggested his judgement might not correspond with the “real” severity of his situation:

> ‘You could say that I, ehm, I probably have a bit different pain tolerance than my wife. To be completely honest I’m a bit more of a wimp. Hehe. And I’ve probably spent more time at the doctor’s than I strictly speaking needed to. And that’s a bit of a shame too because then I take up time that maybe they could have spent on people who were really ill […] I probably experience it as more severe and painful than what it really is.’

Eric seemed to suggest that while he might feel that something is severe, each condition has an objective level of severity, independent of his own assessment.

Severity as objective fact centered on the idea of set criteria and that the severity of a condition depends on whether such criteria are fulfilled. Participants did not agree what such criteria should be. Some volunteered examples such as prognosis and chronicity, and suggested conditions with good prognoses were less severe that those with poor prognoses, or that chronic conditions were more severe than non-chronic ones. There were also suggestions that severity could be considered along some form of scale, where the severity of a condition might be seen to increase the lower one’s age, or the more pain one has, or by the degree of loss of function.

> Sandra (C19): ‘The younger, the more severe a condition should be considered to be.’

> Thomas (C2): ‘Severity is first and foremost the degree of ailment and the duration and the loss of functioning.’

The notion of severity as extrapersonal appeared central to this objective, criteria-based conceptualisation, with emphasis on health outcomes within and across patient groups. This notion often arose in response to subjective conceptualisations of severity. Justin (C14), for instance, argued that individual experience insufficiently described severity, and suggested that applying subjective interpretations in a healthcare setting would be inappropriate:

> 'If we're talking about a definition of severity then those subjective things can't be included. Even if I think that, eh, it might be experienced as severe for some…but if you're going to define it, I don't think that should be included.'

Building on this, several participants suggested that a subjective assessment of severity would also be impractical in a broader healthcare context:

> John (C18): 'It would set some impossible standards for us as a society, if we have to handle every individual's, let's say, 'created crisis'. What you feel as a crisis but that isn't one. And if society has to deal with that then this is hopeless. That won't even be possible.'

> Amy (C10): 'I think we agree that the severity criterion is very difficult to determine from an individual perspective. Because to the individual [their illness] will mean so much either way […] so how on earth would we place ourselves in the minds of these different people to kind of determine how they view the illness they've got?'

Fairness also arose within the objective conceptualisation. Discussing the distribution of vaccines during the SARS-CoV-2 pandemic, some expressed that it was deeply unfair to consider infection in politicians and members of the royal family more severe—and therefore prioritise vaccines to them—than for other members of society. It appeared that, to ensure fairness, severity should be determined by the same objective standards across all individuals based on objective criteria pertaining to diagnoses generally.

John (C18) presented a different argument for ensuring fairness through objective standards, which we did not identify elsewhere in the data. He argued that considerations of how individuals handle their condition would be unfair to patients who adapt well to their illness. He seemed concerned that patients who rehabilitate well would be punished for their efforts by no longer requiring or receiving support from the healthcare system, while those who do not put in the same effort would be rewarded by receiving continued support. When discussing the idea of directing resources to those who had, across a lifespan, a greater health loss than others, and thus differentiating between individuals within patient groups, John (C18) stated:

> ‘Then you’d punish those who have maybe led a good and healthy life and been healthy. He’ll recover and not be as taken care of as much as the other. That can’t be right?’

Severity as situation dependent

A third theme centered on the idea that severity depends on the context surrounding the person. This emphasis on context appeared to be represented by three subthemes, based on individual circumstances, the effects of a condition on those around the person, and the impact of illness at a societal level.

Severity and the social effects of illness

Some expressed that severity was tied to how illness affects relational, social, and work-related circumstances. As such, severity was about how a condition affects the individual’s life in broader terms, such as the ability to parent, to work, and to enjoy hobbies or social activities. An illness or condition thereby appeared to be considered more severe if it affects one’s ability to function with it.

> Carol (C20): ‘What I’m thinking of is if you fall out of working life. Or if you fall out of hobbies you have. Or if you fall out of your social network. Then I think it’s a severe condition. Because you’re no longer, you’re not really part of normal life anymore.’

> Sara (C16): ‘Severity isn’t just about exactly what your situation is, as in exactly which disease you have, exactly what kind of painkiller you’re taking or…it has something to do with how you experience your situation and how you live with your situation.’

The resources surrounding an individual were also considered relevant to severity and how a condition could impact the individual. The support system surrounding a patient was one such resource, and some stated that a condition could be perceived as more severe in the absence of such a network. People's financial situation was another example of how personal resources could modify the severity of a condition.

> Sara (C16): 'Severity maybe depends on what kind of support system you have around you […] it's more severe to be ill if you don't have a stable personal economy, or a lot of people around you to help. That can also affect how severe something is.'

> Melissa (C20): 'It's less severe because she can buy herself help […] So it creates less severity when you're resourceful.'

Severity and the effects of illness on others

When discussing the situational nature of severity, some related this to the effect a condition might have on those around the patient. Examples included a child affected by their parent's illness, a family bereft of a beloved grandmother, or a social group losing a much-loved friend. A condition seemed to be considered more severe if people beyond the patient are affected.

> Sandra (C19): 'When you're considering severity then you can't just see the individual, you need to see everyone around […] when you're considering the one patient you need to think about who is standing around this patient, who will suffer if you don't prioritise it. What will happen to those around them?'

There were also suggestions this could have a cumulative effect, i.e., the more people affected, the more severe the condition. Speaking about illness generally, Matthew (C4) stated that:

> 'I think the more people it affects, the more severe the illness is.'

Parenthood appeared to be considered especially relevant, and illness in a parent could be more severe due to the effect their illness might have on their child. While talking about parenthood, Mary (C18) expressed that the impact of losing a parent is so substantial that a life-threatening condition should be considered more severe for those with children than for those without:

> ‘I have a brother with three kids, I have no kids. I think that it’s more important that he lives than that I live, if you had to choose between us.’

Severity and the effects of illness at a societal level

Severity was also seen as related to the effects of a condition at a societal level, with illness considered more severe if it induced negative effects on society, such as large costs associated with treatment, or a reduction in productivity. The opportunity cost of care was also pointed out as a concern at the societal level: the more resources directed towards the healthcare system, the less is available for other sectors. Emily (C19) suggested that severity also related to the implications of a condition outside the healthcare sector:

> ‘I’m relatively young, and I’m worried about the welfare state in the future, pension schemes, can we afford to treat people, can we afford to develop good enough schools, nurseries, work for everyone […] that one should focus more on prevention and trying to stop illness before it becomes too severe, before you fall out of work, so even fewer of us can contribute to the welfare state […] I think that’s also very important to consider when we’re talking about severity.’

Paul (C5) pointed to the negative impact of mental health problems, beyond the effects on the patient. He argued that conditions like psychosis could be associated with increased crime rates, and that such impacts should also be considered:

> ‘The societal consequences can be enormous […] I’m thinking about the indirect consequences, that those are also part of the severity criterion. Or should be.’

Fairness again came into play, but here it appeared to be associated with the relational and societal burdens of illness. David (C21) expressed concern about what was fair at a societal level:

> ‘I personally think benefit for society should also be considered. Eh, if you help someone who will lead to a greater burden for society that’s like buying yourself a problem. If you help someone who quite frankly contributes to creating goods for the community, then go for it […] My point it that the benefit for society is also important.’

Rather than determining severity in a fair way for the individual as in the subjective theme, or fairly within and across patient groups as in the objective theme, there was an emphasis on determining severity in a manner that is fair to the wider society. As such, it appeared that determining severity fairly entails a consideration of the effects on the wider society. A fair determination of severity, it seemed, should also take into consideration what the cost of illness is at a societal level.

Discussion

The participants provided rich, detailed, and differing descriptions of severity. We identified three interrelated conceptualisations of the term, namely severity as subjective experience, severity as objective fact, and severity as situation dependent. The disparity between these resonates with the ambiguity surrounding severity in priority-setting literature and policy (11,16,17,29), and our results suggest severity is a complex term also for the citizenry. Our discussion starts with this issue of complexity and what it tells us about the difficulty of conceptualising severity. Observing that our three themes touch on central debates in the literature on priority setting generally and severity specifically, we compare our findings to positions within some of these debates. As fairness seemed central to how severity was perceived, we also discuss the different ways fairness arose in our material. Finally, we argue that there is considerable dissonance between public conceptualisations and policy operationalisations of severity.

Severity: an inherently complex term

While 'severity' is a common, everyday term, it also appears to be inherently complex. As participants discussed the multitude of issues they associated with the term in a healthcare context, they connected and contrasted different interpretations and disagreed among each other (and sometimes themselves) on exactly what severity is and how to judge whether a condition is severe or not. This supports previous findings (12). The disparity between the views we uncover demonstrates the difficulty of capturing what severity really is, and suggests that severity is a polyvalent concept (35), with discrepant assumptions and emphases underpinning how it is understood and applied. Furthermore, participants' views on severity were not contained within the three analytic themes (i.e., these represent themes across participants, not individual views). Participants expressed that severity can mean many different things, and the same participant could express views aligned with more than one conceptualisation. Severity thereby appears to be a fluid and dynamic notion. As such, severity appears not only to be polyvalent, but—for our participants, and depending on the topic of discussion—to mean more than one thing at once.

The subjective and objective conceptualisations of severity in many ways represent opposite ends of a continuum. The focus on the subjective and the individual, their experience of their situation, and their unique position to assess its severity is reminiscent of a phenomenological approach to illness (36–38), emphasising the role of subjectivity and how lived experience uniquely informs understandings of illness (38,39). Severity as an objective, criteria-based description of disease, on the other hand, bears comparison to criteria-driven approaches to healthcare (40). These endpoints on the continuum relate to familiar positions within the wider healthcare literature and map onto a longstanding debate on whether health state evaluation should be based on individual, subjective evaluation or objective, generalisable standards (41,42). This debate is particularly current in relation to severity in the field of genomics, where severity appears in laws and policies but remains undefined (43–46).

Within genomics, much like in our data, there is discord on whether severity should be determined by referring to patients' lived experience or objective, criteria-based standards (43,47). However, while both the genomics literature and the healthcare literature more widely tend to treat subjective and objective approaches as distinct, overall our participants appeared ambivalent about the degree to which severity is a subjective or objective concept, or whether it contains elements of both. This reinforces the notion that severity could be seen as existing somewhere along a continuum between the two.

In the situational conceptualisation of severity, emphasis is on the effects of illness and where these effects are located, from considering the social effects of illness on the patient, to the effects of illness on those surrounding the patient, and finally to the effects of illness at a societal level. Some of these effects appear to be located outside the healthcare sector, beyond treatment and care. This is reminiscent of the societal perspective sometimes adopted in health economic evaluation (14,48), where factors such as absence from (paid and unpaid) work and the burden on family members (i.e., health spillovers) are considered relevant to the decision context (49,50). The subthemes we identify within the situational conceptualisation could be said to represent three orders of effects, stretching from a first order concerning the patient and the social effects of illness on them, to a second order of effects of illness on those immediately surrounding the patient, and finally to a third order of the broader effects of illness at a societal level. It is notable that priority-setting frameworks tend to adopt a healthcare perspective when considering cost-effectiveness evidence, and severity is often operationalised as disregarding the effects of illness beyond the patient and their medical needs (4,7). The orders of effects within the situational conceptualisation, however, demonstrate a concern among the citizenry for the relevance of indirect burdens and costs of illness to severity.

Determining severity fairly

In all three analytic themes, the issue of fairness arose as relevant to how, and on what level, severity is determined. In *'severity as subjective experience'*, concern is with fairness for the individual patient. In *'severity as objective fact'*, emphasis is on ensuring a fair determination of severity across all those using health services. In *'severity as situation dependent',* there is a concern for determining severity fairly at a societal level.

Concern for ensuring a fair determination of severity both for the individual and for society also compare to the literature on health economic evaluation, specifically within the literature on QALYs themselves and who ought to evaluate health states (10,14,51,52). QALY estimates (more specifically, the valuation of health states, which are used to estimate QALYs) commonly rely on preferences elicited from member of the public (53) so that, when severity is operationalised via QALYs, a condition's severity is determined by the preferences of members of the public who do not suffer from that condition. A central argument for employing public preferences in QALY estimations is that, because public funds pay for healthcare, it is fair that the public should determine the relative value of different health states (52). This argument is comparable to the way fairness arose in the situational subtheme concerning societal effects, emphasising the importance of determining severity fairly on a broader, societal level. In the subjective conceptualisation, on the other hand, emphasis was on ensuring fairness by determining severity according to the individual's subjective experience. This aligns with a common critique of public preference-based QALY estimation and ties back to the notion of lived experience, i.e., that patients know their condition best, and are therefore best situated to evaluate it (52).

The emphasis within the objective conceptualisation on applying criteria to determine severity fairly within and across patient groups is somewhat aligned with a central motivation behind the QALY, namely to create a standardised approach to classifying health states across patient groups (54). Here, we also identified a notion of fairness centred on determining severity objectively to avoid punishing those who adapt well to their illness. This view, which was expressed by one participant, stands in stark contrast to the egalitarian principle of concern for the worse-off central in health economic literature, asserting that patients who fail to adapt to illness should not be punished for it (55–58). This alternate notion of fairness, which appears to voice concern for the better-off, compares to a libertarian position (59) and represents an antithetical approach to severity and fairness to that in health economic literature and policy.

Conceptual and operational mismatch

Our results touch on many central debates and contested issues in the priority-setting literature generally and on severity and QALYs specifically. This is an interesting finding in itself, demonstrating that these concerns are not exclusive to the domains of policy and academia, but relate to issues members of the public intuitively care about and point to. Reflecting on the issues brought up in the conversations, there are elements across the analytical themes that might be supported by QALY-based operationalisations of severity, such as the relevance of the risk of death, illness prognosis, and quality of life. However, our results touch on a plethora of additional concerns and appear to contain more than QALY shortfall outcomes. For example, we demonstrate views linked to the relevance of lived experience, and determining severity in a way that is fair to the individual. We observe concern for the non-health related effects of illness, as well as for the effects on family members and friends. And we demonstrate severity approached from a societal perspective, including concern for production loss. These additional concerns are not accounted for in the Norwegian, Dutch, or UK absolute and/or proportional QALY shortfall operationalisations of severity, nor represented by the severity levels within the Swedish model (4–7). This supports our claim that severity is a polyvalent concept which is not neatly defined, applied, or contained—neither within current policy operationalisations of the term, nor between the conceptualisations of severity we identify.

Our findings demonstrate poor alignment between operationalisations of severity in policy and conceptualisations among the general public. This mismatch can lead to miscommunications between policymakers and the citizenry, making it difficult for members of the public to understand and support priority-setting decisions involving the term severity. This could lead to complaints from the public on priority-setting outcomes, and policymakers therefore ought to ensure the operationalisation of terms used in everyday parlance corresponds to the everyday meanings of them when such terms are applied in policies. This, we demonstrate, has not been the case with severity. Our findings suggest the citizenry would not consider current policy operationalisations to sufficiently capture illness severity when considered in a healthcare context.

Strengths and limitations

A strength of this study is the breadth of the sample with regards to age, sociodemographic background, and geographical location. Despite considerable efforts, we recruited few participants with minority backgrounds, and there is an overrepresentation of women and individuals with higher education in the sample. The majority also reported to be in good health, on average somewhat better than the Norwegian population more widely (60). It is possible that a broader and larger sample would have yielded additional perspectives. Due to the SARS-CoV-2 pandemic restrictions, the conversations varied in format (online/in-person) and size. This may have influenced the way in which the conversations progressed. The group setting may also have influenced the degree to which participants felt free to express their personal views. We strived to both identify and avoid interpretive biases (61). As preconceptions could influence interpretations, we sought to avoid biases in coding and analysis by having three authors code the same three transcripts and compare approaches. The analytical process was also a collaborative and reflexive process between all authors. The research was conducted in Norway, and results may not be applicable in other contexts, even if severity is used in other jurisdictions as a priority-setting criterion.

Conclusion

When operationalising severity in priority-setting frameworks, policymakers appear to have taken for granted that severity is a concrete and well-defined concept. Our findings suggest this is not the case. Having explored the knowledge gap on what severity means to the public, the three interrelated conceptualisations we identify suggest that severity is a polyvalent concept. Comparing our findings to the literature on priority setting and severity, it is evident that, while there is some overlap with QALY-based operationalisations, severity involves many additional concerns for the citizenry. Our findings provide policymakers with a richer understanding of what severity means to the public and demonstrate that there appears to be considerable dissonance between public conceptualisations and policy operationalisations of severity.

List of abbreviations

QALY	Quality-adjusted life year
SEVPRI	SEVerity and PRIority setting in healthcare
UK	United Kingdom
US	United States

Declarations

Ethics approval and consent to participate

The Regional Committee for Medical and Health Research Ethics deemed the SEVPRI study outside the remit of the Norwegian Health Research Act (ref. no. 186284). Accordingly, ethical approval was granted by the Data Privacy Officer at Akershus University Hospital Trust following a detailed Data Protection Impact Assessment (*PVO. Nos 20_200* and *21_200*). The research for this paper was performed in accordance with relevant ethical guidelines and regulations, including Guidelines for Research Ethics in the Social Sciences and the Humanities (62) and Guidelines for Research Ethics and Scientific Assessment of Qualitative Research Projects in Medicine and Healthcare (63). All names of people and organisations have been deleted or altered. All participants gave their informed written consent via a consent form ahead of the conversations, and their informed oral consent at the outset of the conversations. They were informed they could withdraw at any stage, but no participants availed of this.

Consent for publication

Not applicable.

Availability of data and materials

The dataset generated and analysed during the current study are not publicly available due to privacy concerns for the participants, but are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This study, under the SEVPRI research project, is funded by the Research Council of Norway (ref. no. 303724).

Author contributions

DGTW contributed to the conception and design of SEVPRI and this study. MSS contributed to the conversation guide, conducted the group interviews, and transcribed the conversations, with the aid of project collaborators (see 'Acknowledgements'). Analysis was conducted by MSS and supervised by JR, DGTW, and HL. All authors contributed to the final manuscript.

Acknowledgements

We wish to thank SEVPRI's Principal Investigator Mathias Barra for valuable input on this manuscript, and for his and project collaborator Borgar Jølstad's contributions to group interview facilitation and transcription. We also wish to thank Tone Breines Simonsen (TBS) for her contribution to participant recruitment, and for transcribing one of the group interviews.

References

1. Hirose I, Bognar G. The ethics of health care rationing: an introduction. Routledge; 2014.

2. Mitton C, Donaldson C. Health care priority setting: principles, practice and challenges. Cost Eff Resour Alloc CE. 2004 Apr 22;2:3.

3. Ahumada-Canale A, Jeet V, Bilgrami A, Seil E, Gu Y, Cutler H. Barriers and facilitators to implementing priority setting and resource allocation tools in hospital decisions: A systematic review. Soc Sci Med. 2023 Apr 1;322:115790.

4. Magnussen J, Aaserud M, Granaas T, Magelssen M, Syse A, Celius EG, et al. På ramme alvor - Alvorlighet og prioritering. Department of Health; 2015 Oct p. 71.

5. Riksdagsförvaltningen. Prioriteringar inom hälso- och sjukvården Socialutskottets Betänkande 1996/97:SoU14 - Riksdagen [Internet]. 1997 [cited 2022 Nov 29]. Available from: https://www.riksdagen.se/sv/dokument-lagar/arende/betankande/prioriteringar-inom-halso--och-sjukvarden_GK01SoU14

6. Schurer M, Matthijsse SM, Vossen CY, van Keep M, Horscroft J, Chapman AM, et al. Varying Willingness to Pay Based on Severity of Illness: Impact on Health Technology Assessment Outcomes of Inpatient and Outpatient Drug Therapies in The Netherlands. Value Health. 2022 Jan;25(1):91–103.

7. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. 2022;181.

8. Norheim OF. Priority To The Young Or To Those With Least Lifetime Health? Am J Bioeth. 2010 Apr 9;10(4):60–1.

9. Ministry of Health and Care Services. Meld. St. 38 (2020-2021) Nytte, ressurs og alvorlighet— Prioritering i helse- og omsorgstjenesten [Internet]. Helse- og omsorgsdepartmentet; 2020 [cited 2023 May 14]. Available from: https://www.regjeringen.no/contentassets/85eb048de42342e4b789913cb8b1b218/no/pdfs/stm202020210038000dddpdfs.pdf

10. Weinstein MC, Torrance G, McGuire A. QALYs: The Basics. Value Health. 2009 Mar 1;12:S5–9.

11. Olsen JA. Hva menes med «sykdommens alvorlighetsgrad»? Tidsskr Den Nor Legeforening [Internet]. 2013 Jan 8 [cited 2022 Jun 16]; Available from: https://tidsskriftet.no/2013/01/kronikk/hva-menes-med-sykdommens-alvorlighetsgrad

12. 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 Jun 1;122(6):630–7.

13. Rolstad K. Prioritering i helsevesenet: Verdier, konfliktlinjer og veivalg: en analyse av prioriteringsutredningene i Norge og Sverige. [Göteborg]; 2003.

14. Drummond M, Brixner D, Gold M, Kind P, McGuire A, Nord E. Toward a Consensus on the QALY. Value Health. 2009 Mar;12:S31–5.

15. Nord E. Cost-Value Analysis in Health Care: Making Sense Out of QALYS. Cambridge University Press; 1999. 188 p.

16. Barra M, Broqvist M, Gustavsson E, Henriksson M, Juth N, Sandman L, et al. Severity as a Priority Setting Criterion: Setting a Challenging Research Agenda. Health Care Anal. 2019 May 22;1–20.

17. 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 Jun 1;116(2):281–8.

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

19. McHugh N. Eliciting public values on health inequalities: missing evidence for policy windows? Evid Policy. 2022 Nov;18(4):733–45.

20. Dolan P. The measurement of individual utility and social welfare. J Health Econ. 1998 Jan 1;17(1):39–52.

21. Gyrd-Hansen D. Investigating the social value of health changes. J Health Econ. 2004 Nov;23(6):1101–16.

22. Nord E. The trade-off between severity of illness and treatment effect in cost-value analysis of health care. Health Policy. 1993 Aug 1;24(3):227–38.

23. Ryynänen OP, Myllykangas M, Kinnunen J, Takala J. Attitudes to health care prioritisation methods and criteria among nurses, doctors, politicians and the general public. Soc Sci Med 1982. 1999 Dec;49(11):1529–39.

24. Ubel PA, Loewenstein G, Scanlon D, Kamlet M. Individual utilities are inconsistent with rationing choices: A partial explanation of why Oregon's cost-effectiveness list failed. Med Decis Mak Int J Soc Med Decis Mak. 1996 Jun;16(2):108–16.

25. Green C. Investigating public preferences on 'severity of health' as a relevant condition for setting healthcare priorities. Soc Sci Med. 2009 Jun;68(12):2247–55.

26. Oddsson K. Assessing attitude towards prioritizing in healthcare in Iceland. Health Policy Amst Neth. 2003 Nov;66(2):135–46.

27. 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 Jul 1;18(6):671–83.

28. Abelson J, Lomas J, Eyles J, Birch S, Veenstra G. Does the community want devolved authority? Results of deliberative polling in Ontario. CMAJ Can Med Assoc J J Assoc Medicale Can. 1995 Aug 15;153(4):403–12.

29. Hausman D. The significance of 'severity'. J Med Ethics. 2019 Aug;45(8):545–51.

30. Stenmarck MS, Jølstad B, Baker R, Whitehurst DGT, Barra M. A severely fragmented concept: Uncovering citizens' subjective accounts of severity of illness. Soc Sci Med. 2023 Aug 1;330:116046.

31. Malterud K. Qualitative Metasynthesis: A Research Method for Medicine and Health Sciences. 1st edition. Milton Park, Abingdon, Oxon ; New York, NY: Routledge; 2019. 128 p.

32. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant. 2018;52(4):1893–907.

33. Barbu CM. Zoom: A spatial data visualization tool. 2014.

34. Crabtree BF. Doing Qualitative Research. SAGE; 1999. 434 p.

35. Pilgrim D. `Recovery' and current mental health policy. Chronic Illn. 2008 Dec 1;4(4):295–304.

36. Hummel RP. Phenomenology in Planning, Policy & Administration. Polity. 1982 Dec;15(2):305–14.

37. Smith DW. Phenomenology. In: Zalta EN, editor. The Stanford Encyclopedia of Philosophy [Internet]. Summer 2018. Metaphysics Research Lab, Stanford University; 2018 [cited 2023 Jun 4]. Available from: https://plato.stanford.edu/archives/sum2018/entries/phenomenology/

38. Toombs SK. The lived experience of disability. Hum Stud. 1995 Jan 1;18(1):9–23.

39. Reynolds JM. The Life Worth Living. Minnesota: University of Minnesota Press; 2022.

40. Marsh K, Lanitis T, Neasham D, Orfanos P, Caro J. Assessing the Value of Healthcare Interventions Using Multi-Criteria Decision Analysis: A Review of the Literature. PharmacoEconomics. 2014 Apr 1;32(4):345–65.

41. Cleary PD. Subjective and Objective Measures of Health: Which is Better When? J Health Serv Res Policy. 1997 Jan 1;2(1):3–4.

42. Mobbs RJ. From the Subjective to the Objective era of outcomes analysis: how the tools we use to measure outcomes must change to be reflective of the pathologies we treat in spinal surgery. J Spine Surg Hong Kong. 2021 Sep;7(3):456–7.

43. Boardman FK, Clark CC. What is a 'serious' genetic condition? The perceptions of people living with genetic conditions. Eur J Hum Genet. 2022 Feb;30(2):160–9.

44. Kleiderman E, Ravitsky V, Knoppers BM. The 'serious' factor in germline modification. J Med Ethics. 2019 Aug;45(8):508–13.

45. Kleiderman E, Ravitsky V, Knoppers BM. 'Serious' factor—a relevant starting point for further debate: a response. J Med Ethics. 2020 Feb;46(2):153–5.

46. Newson AJ, Dive L. Taking seriousness seriously in genomic health. Eur J Hum Genet. 2021 Nov 15;1–2.

47. Wertz DC, Knoppers BM. Serious genetic disorders: Can or should they be defined?*: Defining Serious Genetic Disorders. Am J Med Genet. 2002 Feb 15;108(1):29–35.

48. Drost RMWA, Putten IM van der, Ruwaard D, Evers SMAA, Paulus ATG. Conceptualisations of the societal perspective within economic evaluations: a systematic review. Int J Technol Assess Health Care. 2017 ed;33(2):251–60.

49. Al-Janabi H, van Exel J, Brouwer W, Coast J. A Framework for Including Family Health Spillovers in Economic Evaluation. Med Decis Making. 2016 Feb 1;36(2):176–86.

50. Boccuzzi SJ. Indirect Health Care Costs. In: Weintraub WS, editor. Cardiovascular Health Care Economics [Internet]. Totowa, NJ: Humana Press; 2003 [cited 2023 May 10]. p. 63–79. (Contemporary Cardiology). Available from: https://doi.org/10.1007/978-1-59259-398-9_5

51. Edgar AR, Salek S, Shickle D, Cohen D. The ethical QALY: ethical Issues in healthcare resource allocations [Internet]. Haslemere, UK: Euromed Communications; 1998 [cited 2023 Jul 1]. 168 p. Available from: https://orca.cardiff.ac.uk/id/eprint/46019/

52. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. Br Med Bull. 2010 Dec 1;96(1):5–21.

53. Stamuli E. Health outcomes in economic evaluation: who should value health? Br Med Bull. 2011 Mar 1;97(1):197–210.

54. Spencer A, Rivero-Arias O, Wong R, Tsuchiya A, Bleichrodt H, Edwards RT, et al. The QALY at 50: One story many voices. Soc Sci Med. 2022 Mar 1;296:114653.

55. Brock D. Priority to the worse off in healthcare resource prioritisation. In: Rhodes R, Battin MP, Silvers A, editors. Medicine and Social Justice: Essays on the Distribution of Health Care. Oxford University Press; 2002.

56. Nord E. Concerns for the worse off: fair innings versus severity. Soc Sci Med. 2005 Jan;60(2):257–63.

57. Norheim OF, Emanuel EJ, Millum J, editors. Global Health Priority-Setting: Beyond Cost-Effectiveness [Internet]. New York: Oxford University Press; 2019 [cited 2022 Jun 16]. 352 p. Available from: https://oxford.universitypressscholarship.com/10.1093/oso/9780190912765.001.0001/oso-9780190912765

58. Ottersen T, Mæstad O, Norheim O. Lifetime QALY prioritarianism in priority setting: quantification of the inherent trade-off. Cost Eff Resour Alloc. 2014;12(1):2.

59. Grossman LA. The Origins of American Health Libertarianism. Yale J Health Policy Law Ethics. 2013;13:76.

60. Garratt AM, Hansen TM, Augestad LA, Rand K, Stavem K. Norwegian population norms for the EQ-5D-5L: results from a general population survey. Qual Life Res. 2022 Feb 1;31(2):517–26.

61. Guest G, M.MacQueen K, E.Namey E. Applied Thematic Analysis [Internet]. SAGE Publications, Inc.; 2012 [cited 2023 Mar 20]. Available from: https://methods.sagepub.com/book/applied-thematic-analysis

62. NESH. Forskningsetikk. 2021 [cited 2023 Aug 31]. Guidelines for Research Ethics in the Social Sciences and the Humanities. Available from: https://www.forskningsetikk.no/en/guidelines/social-sciences-humanities-law-and-theology/guidelines-for-research-ethics-in-the-social-sciences-humanities-law-and-theology/

63. NEM. Forskningsetikk. 2009 [cited 2023 Aug 31]. Guidelines for Research Ethics and Scientific Assessment of Qualitative Research Projects in Medicine and Healthcare. Available from: https://www.forskningsetikk.no/retningslinjer/med-helse/vurdering-av-kvalitative-forskningsprosjekt-innen-medisin-og-helsefag/

Social Science & Medicine 330 (2023) 116046

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

A severely fragmented concept: Uncovering citizens' subjective accounts of severity of illness

Mille Sofie Stenmarck [a,b,*], Borgar Jølstad [a], Rachel Baker [c], David GT. Whitehurst [d], Mathias Barra [a]

[a] The Health Services Research Unit – HØKH, Akershus University Hospital HF, Lørenskog, Norway
[b] Faculty of Medicine, University of Oslo, Norway
[c] Yunus Centre for Social Business and Health, Glasgow Caledonian University, Scotland, UK
[d] Faculty of Health Sciences, Simon Fraser University, Canada

ARTICLE INFO

Handling Editor: Joanna Coast

Keywords:
Healthcare
Resource allocation
Severity
Priority setting
Q-methodology

ABSTRACT

Universal healthcare is constrained by national governments' finite health resources. This gives rise to complex priority-setting dilemmas. In several universal healthcare systems, the notion of severity (Norwegian: 'alvorlighet') is a key consideration in priority setting, such that treatments for 'severe' illness may be prioritised even when evidence suggests it would not be as cost-effective as treatment options for other conditions. However, severity is a poorly-defined concept, and there is no consensus on what severity means in the context of healthcare provision – whether viewed from public, academic, or professional perspectives. Though several public preference-elicitation studies demonstrate that severity is considered relevant in healthcare resource distribution, there is a paucity of research on public perceptions on the actual *meaning* of severity. We conducted a Q-methodology study between February 2021 and March 2022 to investigate views on severity amongst general public participants in Norway. Group interviews ($n = 59$) were conducted to gather statements for the Q-sort ranking exercises ($n = 34$). Data were analysed using by-person factor analysis to identify patterns in the statement rankings. We present a rich picture of perspectives on the term 'severity', and identify four different, partly conflicting, views on severity in the Norwegian population, with few areas of consensus. We argue that policymakers ought to be made aware of these differing perspectives on severity, and that there is need for further research on the prevalence of these views and on how they are distributed within populations.

1. Introduction

National healthcare systems operate with finite budgets, and priority-setting decisions are difficult and unavoidable (Hirose and Bognar, 2014). In a number of healthcare systems, severity is a key concept in priority setting. Treatments for 'severe' conditions may be prioritised even when evidence suggests there might be more cost-effective investments available for other conditions (Barra et al., 2019). Yet, severity is a multifaceted and arguably poorly-defined concept, and there is no consensus on what severity means in the context of healthcare provision – whether viewed from public, academic, or professional perspectives. In this paper, we contribute towards a more nuanced discussion of the meanings attached to 'severity' amongst the public in Norway.

1.1. *Severity in priority-setting policy*

Most jurisdictions with universal healthcare systems place value on and stipulate requirements for public consultation and transparency. This is also true of Norway (Norheim et al., 2019). Four separate government-appointed commissions have produced green papers (NOU 1987:23, NOU, 1997:18, NOU, 2014:12; NOU, 2018:16) that develop and establish statutory principles for priority setting in healthcare. Three priority-setting criteria have emerged: the first two relate to *cost-effectiveness;* the latter is *severity*. The purpose of the severity criterion is to allow for a higher priority than cost-effectiveness suggests for treatment options targeting conditions that are particularly severe. The severity criterion thus modifies standard decision rules used in cost-effectiveness analyses (Norheim, 2010).

In Norway, like many other jurisdictions, for the purpose of cost-

---

[*] Corresponding author. HØKH, Akershus University Hospital, Sykehusveien 25, 1468 Nordbyhagen, Norway.
E-mail address: msstenmarck@gmail.com (M.S. Stenmarck).

https://doi.org/10.1016/j.socscimed.2023.116046
Received 24 October 2022; Received in revised form 9 June 2023; Accepted 21 June 2023

effectiveness assessments, the *quality-adjusted life year (QALY)* is the *de facto* measure of outcome in economic evaluation (Canadian Agency for Drugs and Technologies in Health, 2017; National Institute for Health and Care Excellence, 2022; Norheim et al., 2014). The QALY combines quality and quantity of life in a single outcome (Weinstein et al., 2009) and the cost-effectiveness of treatment options are given as cost-per-QALY estimates (Drummond et al., 2009). Severity criteria for priority setting are used in several European countries, including the Netherlands (Schurer et al., 2022), Sweden (Riksdagsförvaltningen, 2018), and England (National Institute for Health and Care Excellence, 2022). In Norway, a QALY-based operationalisation of severity – *absolute QALY shortfall* (AQS) (Magnussen et al., 2015) – has been adopted; NICE's approach combines AQS *and proportional QALY shortfall* (PQS); PQS is used in the Netherlands (Reckers-Droog et al., 2018). Sweden operates with a more qualitative approach to severity (Riksdagsförvaltningen, 2018).

1.2. Severity defined

The notion that health resource allocation should incorporate a concern for severity is often grounded in ethical theories of distributive justice, emphasising the claims of the worse off (Daniels, 1985; Rawls, 1999). Attempts at defining severity have focused largely on measurable interpretations of the term, based on QALYs (Drummond et al., 2009; Lakdawalla and Phelps, 2020; Nord, 1999). However, severity remains a controversial concept (Hausman, 2019): there is ambiguity in both policy and academic literature, where the term severity is often invoked but lacks a widely accepted definition (Barra et al., 2019; Stein et al., 1987). For example, a questionnaire distributed amongst Norwegian healthcare workers, decision-makers, and patient organisations found widely differing views on the meaning of severity and no unambiguous understanding of how it should be applied (Magnussen et al., 2015).

1.3. Public views on severity

Eliciting public views is central for policy decisions (Tenbensel, 2010) and increases the democratic legitimacy of those policies (Rutgers, 2015). There is an increasing interest in involving the public in healthcare priority setting (Mullen, 1999), and several public preference-elicitation studies have attempted to establish preferences for prioritising the severely ill (Diederich et al., 2012; Dolan and Shaw, 2003; Green, 2009; Gu et al., 2015; Linley and Hughes, 2013; Nord and Johansen, 2014; Skedgel et al., 2022). A systematic review of empirical studies of public preferences on severity in the context of health evaluation (Shah, 2009) captures studies ranging from small-sample focus groups (including Abelson et al., 1995; Dolan, 1998; Dolan and Cookson, 2000; NICE Citizens Council, 2008) to large population surveys (including Gyrd-Hansen, 2004; Oddsson, 2003; Ubel, 1999), with general population, healthcare worker, and researcher participants. The review illustrates the lack of coherence on severity and the studies use different definitions (e.g. based on QALYs; in terms of 'need'; or related to 'worse-off-ness'). Some studies do not provide any definition of severity, and some do not mention the term severity during data collection, but appear to assume that when asked about e.g. trade-offs between groups according to need, participants align 'need' with 'severity'. There is no sound foundation for the assumptions that (1) participants from these studies uniformly or conceptually associate greater QALY-losses with more severe conditions, or (2) that such a notion of severity is the same that features in a call for higher priority for the severely ill.

We know from empirical literature that severity matters, but we do not know *what the public take severity to mean*: many preference-elicitation studies explore the degree to which (a particular definition of) severity is valued, but not what participants mean by 'severe'. This is problematic because the term *severity* arguably invokes concepts from various domains. Severity can be associated with a poor prognosis, a high risk of fatal outcome, the degree of suffering, urgency, the burdens placed on family members, the magnitude of the estimated health loss (Olsen, 2013; Wittenberg and Prosser, 2013), or as moral impetus to act (Solberg et al., 2023). AQS, for example, might represent 'severity' for some, or it might be overlapping with severity. It might also be the case that participants in these studies are in favour of increasing priority for patients with a high AQS, but for other reasons than 'severity'.

The motivation behind this study is to investigate subjective views on the meaning of severity to paint a rich and detailed picture of accounts of severity. To this end, we use Q-methodology (Brown, 1993; Watts and Stenner, 2022) to study these accounts and locate shared viewpoints.

2. Materials and methods

Q-methodology combines qualitative techniques with exploratory factor analysis to study subjectivity. It is a well-established method used to identify and describe *shared* viewpoints on a topic, areas of consensus, and distinctions between viewpoints. We direct the interested reader towards seminal and introductory texts (Baker et al., 2014; Brown, 1993; Damar and Sali, 2022; McKeown and Thomas, 2013; Stephenson, 1935; N. van Exel and de Graaf, 2005; Watts and Stenner, 2022), as well as a vast array of applied studies using Q in a range of fields (Cuppen et al., 2010; McHugh et al., 2019). Our study was conducted in three stages. In Stage 1 we developed a *Q-set* of statements about the meaning of severity through facilitated group interviews with general population participants. In a second stage of *Q-sort*-exercises, participants sorted the Q-set of statements according to how much they agreed or disagreed with each statement. The third stage used factor analysis to identify similar patterns of Q-sorts and to interpret those factors.

2.1. Stage 1: deriving the Q-set

A Q-set represents the 'universe of opinions' on the topic of study (van Exel et al., 2015). There are several ways to develop a Q-set, either from existing sources or by generating statements through interviews. We conducted group interviews to generate statements for the Q-set. Ahead of the group sessions, a brief introduction to the topic of severity was prepared by authors MSS, BJ and MB (Supplementary material A). This introduction, as well as facilitation of the interviews, was piloted with a user panel at Akershus University Hospital consisting of eight members of the public across different demographic backgrounds. Due to SARS-CoV-2 pandemic restrictions, data collection commenced online using Zoom (Barbu, 2014), before converting to physical groups once restrictions were lifted.

Purposive sampling (Malterud, 2019) aimed to elicit views from participants expected to have differing perspectives. We hypothesised that age, education level, socioeconomic background, health status, and geographical region were relevant characteristics. We monitored variation across these characteristics throughout the recruitment process and sought to fill any gaps by seeking out individuals with underrepresented characteristics. We recruited through the project's social media platforms (Facebook and Twitter) by asking our professional and personal networks (snowballing) to disseminate a link to an online recruitment platform on the *Nettskjema* infrastructure hosted by the University of Oslo. Posters advertising the project were hung in public spaces and GP's waiting rooms in two large Norwegian cities (Oslo and Bergen). Recruitment lasted from February to July 2021, and group interviews were conducted February to May (online) and May to July (physical). We aimed for approximately 60 participants to reach saturation (Malterud, 2019). Inclusion criteria were Scandinavian-speaking adults (age $\geq$18). We determined saturation according to data redundancy, i.e. once participant viewpoints became repetitive (Saunders et al., 2018) and no new Q-statements were generated.

Group interviews were facilitated by one to three authors (MSS, BJ, MB). Participants submitted informed, written consent, and were subsequently introduced to the topic of severity during a brief presentation, developed by authors MSS, BJ, and MB (Supplementary material A). As a note, direct translation of severity from Norwegian ('alvorlighet') to English is difficult, as it is a more all-encompassing term in Norwegian. We use the term severity in relation to 'illness', where we take illness to cover different descriptors of ill health, such as injury, illness, sickness, and disease.

A facilitated group discussion explored participants' views of severity, designed to allow discussions to develop uninterrupted, focused on perspectives participants brought up. Discussion was supplemented with topics from a pre-prepared list of possible aspects of severity, to prompt participants to discuss certain topics. The list was compiled following a comprehensive search of the literature (Barra et al., 2019), and was updated dynamically throughout the interview period as participants raised new issues (Supplementary material A). Finally, participants completed a questionnaire about socioeconomic status, health status, and situations that may have affected their views on severity (Supplementary material B). Participants in physical groups filled out the questionnaires on paper; online participants were contacted by a facilitator to complete the questionnaire by telephone.

All group interviews were audio-recorded and transcribed (in Norwegian) by MSS and BJ. The same authors coded the transcripts by highlighting (without judgement) every statement that expressed a view about what severity means. These candidate statements for the Q-set were examined for intelligibility and relevance, with the aim of achieving coverage of all the identified issues (Watts and Stenner, 2022). Duplicate statements were removed, very similar statements were merged, and those expressing several opinions were broken down (Baker et al., 2017). All statements were categorised using group and participant codes to ensure breadth in selected statements. The statement set was reviewed at multiple stages and discussed by co-authors to reach a final Q-set, with a view to representing the breadth of opinion expressed across the group interviews. The Q-set was piloted with a convenience sample of 14 colleagues, who are academics across different field, aiming to ensure the statements and instructions (Supplementary material C) were easy to understand. Pilot participants responded positively to the Q-set and did not identify any missing statements.

2.2. Stage 2: Q-sort exercise

The Q-sort is the main source of data in a Q-study. Participants are asked to rank statements onto a grid according to a standard instruction. A new set of participants were recruited for the Q-sort exercises, in the same manner as in Stage 1 (January to March 2022), seeking breadth of demographic profiles and recruiting in two locations (Oslo and Bergen). The exercises were conducted face-to-face. Participants were gathered in groups and introduced to the study and instructions together, then completed the Q-sort independently. They were presented with the Q-set (with each statement represented on an individual card) and a Q-sort grid (Fig. 1). Participants were first asked to sort all cards into three piles: 'agree', 'disagree', and 'neutral'; then to sort the cards onto the grid, with the placement of each card reflecting the degree of agreement, from most agree (+6) to most disagree (−6) (Supplementary material C). The grid forces distribution of cards into a particular number in each column, with one square for each card. This is standard in Q-methodology (Watts and Stenner, 2022) and although a forced (compared to free) distribution could appear to restrict participants' viewpoints, comparisons of the two indicate a negligible effect (Brown, 1993; Hess and Hink, 1959). Finally, participants were asked to elaborate on their rankings on a separate form (Supplementary material D) and answer questions about their sociodemographic status (using the same questionnaire as in Stage 1).

2.3. Stage 3: exploratory factor analysis and interpretation of Q-sorts

In Q-methodology, the 'best' factor solution is determined by the qualitative interpretability of each solution, as well as its statistical qualities (McKeown and Thomas, 2013). It is customary to fit models with varying numbers of factors and judge their merits according to qualitative readings of the resulting factor arrays, rather than purely on statistical grounds.

Data analysis was conducted using dedicated Q-methodology software (*KenQ-Analysis Desktop Edition (KADE)* version 1.2.1) (Banasick, 2019). Correlations between Q-sorts were calculated and clusters of participants identified using by-person factor analysis. KADE fits a centroid factor analysis (with Horst 5.5 centroid factors) with varimax rotation to identify shared viewpoints among the participants (Watts and Stenner, 2022). The resulting factors are represented by a factor array, or composite Q-sort, for each factor based on the Q-sorts of the factor exemplars (participants with high, pure loading on one factor only). These composite sorts represent a weighted (according to factor loadings) compromise between the Q-sorts of participants flagged as exemplars for that factor (Watts and Stenner, 2022) and form the basis for interpretation.

In determining a preferred factor solution, we considered the number of Q-sorts flagged for each factor to ensure that factors represent shared perspectives between Q-sorters – ideally at least four flagged Q-sorts to be satisfied the factor represents a shared viewpoint (Watts and Stenner, 2022). We considered the correlation between different factor solutions to see whether distinct new factors were identified in solutions with increasing numbers of factors (Watts and Stenner, 2022). Four authors

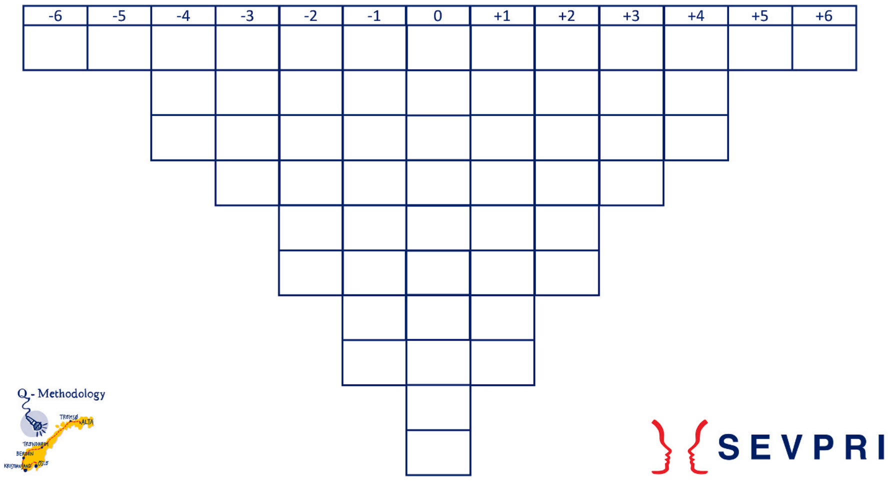


Fig. 1. The grid for the Q-sorting exercise.

(MSS, BJ, RB, MB) examined and produced initial factor interpretations based on the composite Q-sorts and qualitative information for each factor. All authors participated in the deliberative process of factor analysis and discussed interpretations in two separate workshops. All statements in the composite Q-sorts were considered, with particular focus on the salient statements (at either side of the grid), distinguishing statements (placed substantially differently between factors), and consensus statements (placed similarly across factors) (Watts and Stenner, 2022). Non-salient and neutral statements also contributed towards interpretation. Written responses (where participants elaborated on their rankings) uniquely associated with each factor were examined to enhance interpretation and ensure a rich understanding of participants' views.

2.4. Ethics

The Regional Committee for Medical and Health Research Ethics (Regional Ethics Committee South-East B) advised that SEVPRI falls outside their mandate (i.e., the aims and objectives are not regulated by the Health Research Act (Helseforskningsloven, 2008)). Consequently, the Data Privacy Officer at Akershus University Hospital evaluated the project's protocol and advised that the study could be conducted (*PVO. Nos 20_200 and 21_200*). Akershus University Hospital and the Principal Investigator (MB) are responsible for project oversight, including all aspects of ethical research conduct and data privacy.

3. Results

3.1. Stage 1 and stage 2

59 individuals at five locations (Oslo, Bergen, Trondheim, Tromsø, Alta) participated in Stage 1. There were 14 online groups (three with three participants; ten with two participants; and one individual interview) and seven physical groups (one with six participants; two with five participants; two with four participants; one with three participants; and one with two participants), lasting two to 3 hours.

The interview transcripts contained 450 statements on severity. Following coding and statement extraction as outlined above, 53 statements remained, representing distinctive opinions about severity. Three statements were added by the authors, representing viewpoints considered to have theoretical relevance that had not been expressed by participants (statements #54, #55, #56) (Barra et al., 2019). This resulted in a final Q-set of 56 statements.

Thirty-four participants completed the Q-sort exercise in Stage 2, across two locations (Oslo and Bergen). The characteristics of participants involved in the group interviews (Stage 1) and Q-sort exercises (Stage 2) are summarised in Table 1.

Table 1
Participant characteristics.

Age [a]	18–30	31–50	51–66	67+	NA
Q-interview	9	13	24	11	2
Q-sorters	16	9	7	2	–
Sex	**Female**		**Male**	**Other/prefer not to say**	
Q-interview	38 (64%)		19 (32%)	2 (3%)	
Q-sorters	24 (71%)		9 (26%)	1 (3%)	
Religious sentiment	**Religious and/or spiritual Active in a congregation**		**Religious and/or spiritual Not Active in a congregation**	**Neither Religious nor spiritual**	**NA**
Q-interview	11 (19%)		14 (24%)	33 (56%)	1 (2%)
Q-sorters	2 (6%)		4 (12%)	25 (74%)	3 (9%)
Highest education level	**Elementary/ upper secondary**		**BSc/Fagbrev**	**MSc/PhD**	**NA**
Q-interview	9 (15%)		21 (36%)	27 (46%)	2 (2%)
Q-sorters	8 (24%)		16 (47%)	10 (29%)	–
Experience of severe illness [b]	**Transient**		**Chronic**	**Terminal**	**NA**
Q-interview	18 (31%)		30 (51%)	42 (71%)	
Q-sorters	18 (53%)		20 (59%)	23 (68%)	–
Self-reported health	**Very good/ Good**		**Just fine**	**Bad/Very bad**	**NA**
Q-interview	37 (63%)		15 (25%)	6 (10%)	1 (2%)
Q-sorters	19 (56%)		13 (38%)	1 (3%)	1 (3%)

Notes: All percentages are rounded and might not sum to 100%. [a]Age was given in one of the listed age-brackets. [b] Personally or close acquaintance; categories not mutually exclusive. 'Terminal' relevant to acquaintance only.

3.2. Stage 3: exploratory factor analysis and Q-sort interpretation

We computed and inspected factor solutions ranging from two to six factors and selected a preferred factor solution based on the interpretability of the factors as well as their statistical features. We attempted interpretation of the factors (by examining the composite Q-sorts) in the three-, four-, and five-factor solutions. With fewer than three factors, viewpoints were difficult to interpret. With more than five factors, the viewpoints were no longer distinct. We report the four-factor solution, where each composite sort had a coherent and interpretable narrative consistent with the written comments by exemplars. Table 2 shows the factor array for each of the four factors; the statement scores in the rightmost columns of Table 2 permit the positioning of the Q-set back onto the grid for each factor.

Table 3 sets out factor loadings for each participant, representing the correlation between their individual Q-sort and each of the factors. In the four-factor solution, at $p < 0.05$ significant loadings are ≥ 0.262 (McKeown and Thomas, 2013, p.51). Q-sorts were flagged by the following criteria: (i) a significant factor loading (≥ 0.26), and (ii) $>$ 50% of the communality corresponding to the factor (the default in KADE). Participant 23 (R23) can be viewed as an exemplar, with a high loading on Factor III and very little in common with the other factors (Table 3). A flagging algorithm is applied to identify Q-sorts that will contribute to the composite Q-sort through which we interpret each factor. This takes into account the significance of loading on each factor, and the communality (the sum of squared loadings) (McKeown and Thomas, 2013).

Below, each factor is described based on the composite Q-sorts and written comments. These are (necessarily) summary descriptions of the factors. Statements are referred to by number # followed by their score for that factor's composite sort (−6 to +6). Distinguishing statements are indicated by * and **.

3.3. Factor I: 'natural lifespan'

The first factor represents a viewpoint focused on 'the natural course of life', reconciled with the idea that life is inherently finite, and that a life with dignity is preferable to longevity. Severity is associated with the loss of quality of life (#14, +3). Mental illness is particularly severe (#8, +4**), possibly associated with the value placed on autonomy and living life on your own terms and with dignity (#31, +3). People associated with this factor tend to view illness in the young as more severe than illness in older people (#18, +5; #33 + 3) who have already had the chance to live a full life. This is consistent with the notion of a natural lifespan and reinforced by the placement of statements pertaining to

Table 2
Factor arrays.

		Statement scores			
		I	II	III	IV
1	In my mind death is the most severe no matter how old you are. Death is death.	−6**	0	0	−1
2	Diseases that make you die before the life-expectancy in Norway, that is severe.	−1	+2	0	0
3	Severity is first and foremost just death.	−5	−6	−2**	−4
4	Loss of identity is severe. The stronger the loss of identity, the more severe we can say that an illness is.	+2	+1	−1	+1
5	I think it is severe when it's something to do with the brain, how you come across, what you think and stuff like that.	+1	0	+2	+2
6	For someone who has already had lots of illness perhaps more illness won't be a shock, and therefore not more severe than someone who's been well their whole life. So in my opinion illness is more severe for someone who was healthy before.	−3	−3	−3	−2
7	If life can have a good ending, then death feels less severe.	+4	+3	0	0
8	I think mental illness is very severe.	+4**	+1	−1	0
9	When I think of severe, I think of change.	−2	−3*	0*	−2
10	The greater the loss of function, the more severe the illness is.	+1	−1**	+1	+4**
11	Imagine if Arve Tellefsen [famous Norwegian violinist] lost his grip and can't play his violin anymore. That would be much more severe than if I lost a finger. I could still do almost exactly the same as I do now.	**−1**	**0**	**+1**	**0**
12	When you cannot take care of yourself, I think that's severe.	+2	0**	+4	+3
13	I think severity is if you fall out of professional life.	−1	−4**	+3**	−1
14	I think severity is defined from the sense of loss of quality of life.	+3	+1	0	+6*
15	It would be very severe for me with illness that prevents me from living the life I want to live.	+4	−1	+1	+5
16	That there are things that are more severe than dying, that I'm sure of.	+6**	0*	+1	+1
17	The younger, the more severe an illness must be considered to be.	0	+2	+2	−3**
18	A cardiac arrest or a stroke in an 18-year-old is much more severe than in a 98-year-old.	+5	+5	+1*	−1*
19	Severity must be defined by the individual. As in what is severe for each individual.	−1**	+3	−4**	+2
20	I think severity is something the professionals within that field should comment on, not the relatives, not the people who are ill.	0*	−2*	+2**	−4**
21	I think severity is mostly how you take it. That an event or a condition to an insufficient degree describes severity. It's not the event in itself that defines severity, but what it does to you and how you experience it.	0**	+3	−4**	+3
22	Many women survive breast cancer, and some get fully back to work and some have lots of side effects after. And I think that she who has lots of side effects has had a more severe disease.	0	+2	0	−1
23	To experience that you are different, that is severe. But if many others get the same illness, then I would say it becomes less severe.	**−2**	**−1**	**−3**	**−3**
24	It's severe if many others get it. So, severity has something to do with the amount, how many get ill from it.	−4	−5	−4	−2**
25	Pain is severe.	−4**	−1	0	0
26	Death is severe for those left behind, but it's not severe for the person who dies, I think.	−1	−1	−1	−2*
27	Illness isn't severe if it passes on its own, something that doesn't need an intervention from the health services.	+1	+3	+2	−1**
28	If you get immediate treatment and have zero ailments after, then it is not a severe illness.	+1	+2	−1	−3**
29	I think that if you handle a disease or a condition badly then it's more severe.	−1	−2	−2	+1**
30	I think an illness needs to have a closeness to us for it to feel severe.	−3	0*	−3	−2
31	I think it's very severe with a life without dignity.	+3	+4	0**	+4
32	Severity is about being taken seriously.	−3	0	−6**	−1
33	I think it becomes much more severe when it's about children.	+3	+4	+4	−1**
34	It is hard to say anything definitive about what severity really is, I think. There is no right answer.	+2*	+6**	+1	+1
35	All illness can be severe, I think.	−4	−3	−3	0**
36	I think severity is about what the consequence is of not treating.	−1	+1	+3	+2
37	I think how an illness affects everyone around you, that is part of the severity criterion, or should be.	0	−1	−1	+1*
38	If you fall out of the hobbies you've had, or out of the social network you've had. Then, I think it is a severe condition.	+1	−4**	−1	+2*
39	Illness becomes less severe for a resourceful person because she can buy help. So there's less severity when you're resourceful.	−2	−4**	−1	0
40	I think the more people it affects, the more severe the illness is.	−1	−3*	0	0
41	You can't say that because you live a successful life then your illness is more severe than if you're a drug addict. You can't begin to weigh these lives up against each other. Because a life is a life and has the same worth no matter how you've lived or have had to live it.	+2	+2	+6	+4
42	I hurts a bit to say, but I think it's more severe if a 35-year-old mother of two gets cancer than if a 35-year-old single woman gets it.	−3	+1**	−2	−5**
43	I think it's more severe that someone with children gets ill than someone without.	−2	+1**	−2	−3*
44	You can perhaps say that illness is a bit less severe if you have relatives who can support you.	+1	0	−2	−1
45	I think it's severe with illness where it just gets worse and worse. And it's certain that it will only get worse.	+1**	+4	+4	+3
46	As long as you have the possibility to get well, I don't think the disease is severe.	−2	+1**	−2	−4
47	Severity is about the threat of permanent damage.	0	−1	+3	+3
48	If it's an illness that needs to be dealt with very urgently, then it's very severe.	+2	0**	+5**	+2
49	It's as severe to get lung cancer whether you've smoked or not, I think.	+2	−1**	+3	+2
50	Immediately I think illness is less severe if it's your fault you have it. If it's your fault it's less severe.	−2	−2	−5*	−6*
51	If you live far away from the hospital, there's an insecurity in that you might not get help. That's severe.	+1	−2**	+1	+1
52	To get sick is less severe if you are close to the hospital.	0	−2	−1	−2
53	Stigma creates a more severe situation for the individual.	0	+2	+1	+1

(*continued on next page*)

Table 2 (*continued*)

		Statement scores			
		I	II	III	IV
54	How serious the loss of function is depends on the situation. For example, it's more severe to be in a wheelchair if you live somewhere that doesn't facilitate for it, than if you live somewhere where you can still easily get around.	0	0	0	0
55	To be kept alive when you're sick and you want to die, that's severe.	+3	+1	+2	0
56	I think an illness becomes more severe if the treatment becomes a big part of your everyday life. If you have to spend a lot of time in hospital, treatment and stuff.	0	−2**	+2	+1

Notes: * statement is distinguishing $p < 0.01$, ** statement is distinguishing $p < 0.05$. Consensus statements for $p < 0.05$ are highlighted in bold.
Translation: the statements presented here were translated from Norwegian to English by the authors, but were presented to participants in their original Norwegian wording.

Table 3
Factor loadings.

	Factors			
	I	II	III	IV
Participant ID				
R1	**0.53**	0.04	−0.13	**0.55***
R2	**0.34**	0.09	0.04	**0.62***
R3	**0.34**	0.08	**0.38**	**0.47**
R4	0.26	0.02	**0.32**	**0.29**
R5	**0.58***	**0.30**	0.23	0.12
R6	**0.47**	0.18	**0.53**	0.19
R7	0.18	−0.04	**0.66***	0.14
R8	**0.60***	**0.26**	0.17	0.03
R9	−0.21	0.17	**0.28**	**0.48***
R10	0.14	−0.09	0.23	**0.75***
R11	0.05	**0.55***	0.02	0.06
R12	−0.01	0.06	**0.41**	**0.65***
R13	**0.51***	−0.05	0.22	**0.29**
R14	0.13	0.23	0.08	**0.64***
R15	0.16	−0.23	0.37	**0.73***
R16	0.22	0.19	0.33	**0.43**
R17	**0.48**	**0.58***	0.07	0.22
R18	**0.36***	0.12	0.08	0.17
R19	**0.55**	−0.19	0.19	**0.50**
R20	0.24	−0.10	**0.43**	**0.49**
R21	**0.30**	0.13	−0.00	**0.69***
R22	**0.58***	**0.30**	**0.35**	−0.01
R23	0.09	−0.00	**0.64***	0.24
R24	**0.28**	0.23	**0.62***	0.23
R25	−0.05	**0.45***	**0.41**	0.11
R26	0.24	**0.31**	**0.39**	0.18
R27	**0.52***	0.09	0.24	0.11
R28	**0.67***	0.17	0.20	**0.30**
R29	0.26	**0.45***	−0.18	0.04
R30	**0.28**	0.08	**0.35**	0.24
R31	**0.40**	−0.02	**0.28**	0.30
R32	0.23	0.08	**0.60***	0.02
R33	0.17	**0.65***	0.21	−0.07
R34	**0.35**	**0.28**	0.11	**0.52***
Eigenvalues	10.24	2.49	1.80	1.32
Variance	30%	7%	5%	4%

Notes: Significant loadings showing in bold. Flagging indicated by *.

age. Death is seen as natural and not central to the meaning of severity. This is clear from the placing of statements #1 (−6**) and #3 (−5) and statements #16 (+6**), #7 (+4), and #55 (+3), all of which are distinguishing statements for this factor. A good end to life makes death less severe (#7, +4), and being kept alive against your wishes is severe (#55, +3). People associated with this factor tend to reject pain as severe (#25, −4**). This is distinguishing for this factor and fits with the natural lifespan perspective: pain is part of life and does not define severity.

3.4. *Factor II: 'severity is subjective'*

A characterising feature of Factor II is the difficulty of defining severity (#34, +6**). Severity of illness is subjective, and degree of severity is determined by the individual and what *they* consider severe (#19; +3). As such, severity cannot have a general definition, and is not about types of illness or the prevalence of illnesses in a population, but depends on subjective self-evaluation (#21, +3).

In keeping with the explicit emphasis on subjectivity, this factor is largely defined by the rejection of more objective measures of severity. Almost all potential measures of severity are rejected: it is not about prevalence of disease, or loss of identity, nor one's ability to work and enjoy hobbies (#38, −4**), quality of life, or risk of death. Severity is linked to loss of dignity (#31, +4), is increased when there is stigma (#53, +2), and lessened if one can have a "good" death (#7, +3). The subjective experience of severity is connected to stage of life, and illness is more severe when it affects young people (#18, +5; #33, +4; #2, +2; #17, +2). Factor II is distinguished from other factors on the question of whether severity is linked to having children (#42, +1**; #43, +1**), with associated statements placed close to the centre of the grid but rejected by all other factors.

3.5. *Factor III: 'objective measures and triage'*

For Factor III severe illnesses are urgent, progressive conditions that require treatment and care (#48, +5**; #12, +4). This viewpoint might be characterised as a 'medical triage perspective', cognisant of how a doctor prioritises emergency patients. People associated with Factor III tend to consider urgency and prognosis – which are more objectively measurable – as important (#48, +5**; #45, +4). Conditions that effect children are more severe, and in the context of this factor this could be related to the importance of age to medical prognosis (#33, +4). Individuals associated with this factor are also the only ones to support the idea that health personnel have a central role in defining severity (#20, +2**). Statements asserting that severity is subjective, or defined by the individual, or that severity relates to loss of dignity, are ranked distinguishingly low (#19, −4**; #31, 0**). A possible interpretation is that individuals in Factor III consider dignity difficult to measure and consequently too subjective to be relevant.

In keeping with the sense that severity is medicalised in Factor III, issues of culpability or worthiness are not relevant: all lives are equal, and severity is not linked to how successful those lives have been (#41, +6; #50, −5*). This corresponds with a fundamental biomedical ethics principle of treating all patients as equals.

3.6. *Factor IV: 'functioning and quality of life'*

In the final factor, the individual's experience is central, as in Factor II. However, whereas severity is subjectively defined *by the individual* in the latter, in Factor IV severity is determined by the effect a condition has *on the individual.* While severity must be considered in relation to the individual, it is not defined by the individual. Severity is determined by the loss of quality of life (#14, 6*) and how an illness affects day-to-day life: whether you can live the life you want (#15, +5), your level of functioning (#10, +4**), and whether you can still enjoy your hobbies (#38, +2*). The ability to take care of yourself (#12, +3) and live with dignity is also important (#31, +4). Factor IV is distinguished from other factors in the rejection of age (#17, −3**; #18 -1*; #33, −1**). There is also a concern for how illness affects the individual, and a stronger rejection of the notion that next-of-kin or parental responsibilities affect severity. This is consistent with a view that people with equal need should be treated equally, regardless of blame, worth, or parenthood (#50, −6*; #41, +4; #42, −5**).

Interestingly, factors II and IV have a strong negative correlation: if you subscribe to Factor II, you are likely to strongly disagree with Factor IV, and vice versa. This is coherent with our interpretation: they are concerned with similar issues but have directly opposing views of them. While people associated with Factor II subscribe to severity as a subjective experience defined by the preferences of the individual, participants who agree with Factor IV are concerned with the individual experience according to measurable dimensions (e.g., functioning, hobbies, autonomy). Though both groups are concerned with age, in Factor IV age is not relevant, whilst in Factor II this is the only relevant objective measure.

4. Discussion

Severity is a central principle in several health jurisdictions, and has been discussed extensively in the academic literature (Drummond et al., 2015; Millum, 2023; Nord and Johansen, 2014; Shah, 2009; Skedgel et al., 2022). Yet, knowledge of how the public perceives this term is lacking (Barra et al., 2019). Previous preference-elicitation studies have focused on presenting participants with choices between prioritising different patient groups according to QALY-losses and -gains (Shah, 2009), but fail to explicitly connect qualitative explorations of severity with operationalisations that might, or might not, align with colloquial interpretation(s) of the term. In this context, our study is the first of its kind: it presents a rich, detailed analysis of public views on severity and contributes to a more nuanced understanding of the term by investigating subjective accounts on the *meaning* of severity. Our findings suggest that participants' views on the meaning of severity are diverse, at times contradictory, and may or may not overlap with any of the operationalisations of severity encountered in the health economics literature.

Our findings present health economists with a conundrum: is it possible to find a unifying operationalisation based on QALYs that represents the different viewpoints? For example, Factor III, focused on objective measures, and Factor IV, emphasising quality of life, align somewhat with an absolute QALY shortfall operationalisation of severity. However, they also contain elements that do not align with absolute QALY shortfall, such as the preference for the young and the role of dignity, respectively. Factor I, centred on the notion of a natural lifespan, and Factor II, viewing severity as entirely subjective, appear less amenable to QALY-based operationalisations of severity. This is not to say the task of operationalising the viewpoints within a QALY model is impossible. It does seem, however, that this task would prove difficult, and that any one operationalisation is unlikely to represent severity in a way that would be recognisable to all citizens.

The four viewpoints we uncovered are diverse and at times contradictory, with few areas of consensus. They load 'severity' with radically different characteristics, ranging from existential questions on what makes life meaningful, to objective, measurable attributes of illness. Our contribution is thus twofold: we have empirically established that the scholarly debates are not mere intellectual sophistry; these accounts of severity coexist in the colloquial uses of the term. Secondly, using Q-methodology, we have based our interpretations on evidence that can be examined and scrutinised. One of the advantages of Q-methodology when applied to complex qualitative questions is that data and analysis can be made explicit, transparent, and reproducible.

The only basis for consensus between the factors is the relevance of three facets of severity: death, age, and dignity. That is, these attributes matter in all the factors, although they matter in different ways. For example, participants loading on different factors disagree on whether death is severe, but each factor demonstrates strong opinions on death in relation to severity. As for age, people in factors II and III share the view that the severity of a condition varies by age, whilst those in Factor IV reject a relation between age and severity. Those in Factor I occupy a middle position, where age is relevant in determining severity in the young (more severe) and the old (less severe), but only in these extremities of age. The notion of dignity also appears central: in the medicalised viewpoint described in Factor III dignity is not central to severity, while those in the other three factors considered loss of dignity to be very severe.

When elaborating on what severity means, both explicitly during interviews and implicitly through Q-sorts, participants conveyed multifaceted accounts – jointly covering almost all related concepts we conjectured would emerge, including death, age, pain, equality, dignity, desert, quality of life, functioning, and hope. It is not the aim of this study to explore each of those terms, and we submit that *these* terms can be construed as ambiguous, under-theorised concepts. Yet these concepts are used by our participants to distil their subjective accounts of severity, and hence 'severity' invokes an abundance of *thick concepts* (Väyrynen, 2021). As with other thick concepts, severity is described by evaluating it, and evaluated by being described (van der Weele, 2021).

Because interpretations of severity lean on so many different concepts, there are many ways in which our findings could be placed in the literature. We could compare the factors to Western vs. Eastern conceptualisations of health (Sayed, 2003; Tsuei, 1978), where factors I and II are perhaps consonant with Western conceptualisations and factors III and IV closer to Eastern. We could assess them against the social and medical model of disease (Barbour, 1997; Engel, 1992), where Factor III certainly belongs in the medical model while factors I, II and IV arguably contain elements of both. One framework that provides a good background for understanding the factors is Twaddle's analytic triad 'disease', 'illness', and 'sickness'. These represent, respectively, 'medical', 'personal', and 'social' accounts of illness (Hofmann, 2002). According to Twaddle's analysis, 'disease' signals a biomedical view on illness and resonates with the focus on triage and objective evaluations in Factor III. Illness "*is a subjectively interpreted undesirable state of health. It consists of subjective feeling states (e.g. pain, weakness), perceptions of the adequacy of their bodily functioning, and/or feelings of competence*" (Twaddle, 1994, p.10), which aligns with Factor II. 'Sickness' is defined as what afflicts a social identity, and on one account the relationship between the ability to function in society given one's health problems. The focus on objective functioning and illness' impact on quality of life in Factor IV can be construed as existing in the intersection between the biomedical (disease) and the social (sickness), and between disease and illness as these terms are employed by Twaddle and elaborated on by Hofmann. Factor I falls less organically into the triad, but considering this Factor's emphasis on the natural life, it arguably aligns with the narrative of 'sickness'.

While the theoretical relevance of these factors is interesting, the distinctiveness of the factors and the heterogeneity between them highlight a more important issue: there is no consensus on what severity means, neither in academia, in policy, or in public conceptions of severity. This points to our most critical finding: when we as a public talk about severity, we are not all talking about the same thing.

4.1. *Strengths and limitations*

Q-methodology enables rich descriptions of subjective viewpoints and is well-suited for our study. We went to lengths to ensure variation in our sample by conducting online and physical group interviews, and recruiting participants in cities and rural areas. However, it is possible that there are views that are not identifiable in our data. Because the Q-sort sessions were conducted in Oslo and Bergen, citizens from these urban areas are over-represented, and viewpoints of inhabitants of rural areas could be missing. Despite our best attempts, there is also an overrepresentation of women in our sample. Finally, without exception our respondents had some personal experience as patient, caregiver, or both. On the other hand, most people do. Interestingly, the topics on the list of attributes of severity covered all topics discussed by the participants, apart from proximity (represented by statement #52). This suggests that the listed topics, based on the literature, were fairly exhaustive.

Our Q-set was the result of a careful and deliberative process. Nonetheless, the selection and wording of statements could have resulted in a different Q-set depending on decisions made by the research team. Although we sought to extract a Q-set that represents the 'conversational possibilities' (Baker et al., 2014), we may have missed some aspects. However, we asked all participants if there were viewpoints they felt were missing and, overall, participants were satisfied with the topics covered. Though the author team is multinational with varied academic backgrounds, we were wary of potential biases affecting out results. We strived to identify our biases (Guest et al., 2012) and had a reflexive approach to the research process (Flick, 2013). We spent time during all stages of the study reflecting on perspectives we could be missing, and made substantial efforts to include participants with non-majority backgrounds. During the processes of coding statements and interpreting findings, we worked independently before comparing results. We also attempted to reduce social desirability bias during interviews by emphasising that there were no right or wrong answers. While it is never possible to eliminate all sources of interpretive bias, these precautions are likely to have helped us avoid several interpretive pitfalls.

This study posed the question 'what does severity mean'. The results might be different, however, if we asked the public 'which views on severity should be applied in priority setting'. Though this fell outside the scope of this study, it is an important question that should be explored in further work. Finally, future research should explore how the four views we have elicited and described are distributed in a representative population survey.

Author contributions

MB conceived of the study, and RB, DGTW, and MB planned the study design. All authors contributed to the analysis with oversight from MB and RB. MSS wrote the manuscript with input from BJ, MB, DGTW, and RB. All authors discussed the results and contributed to the final manuscript.

Funding

SEVPRI is funded by The Research Council of Norway (ref. no. 303724). MSS, BJ, RB DW, and MB are (partly) supported by this grant.

Declaration of competing interest

None of the project collaborators have any conflicts of interest, nor any financial interests in the project.

Data availability

Data will be made available on request.

Acknowledgements

We wish to thank Tone Breines Simonsen (TBS), Hanne Dale Vonen, and Anders Herting (AH) for their contributions to participant recruitment, TBS for transcribing one of the group conversations, and TBS and AH for assistance with other administrative tasks. We are also grateful for the feedback received from participants at the Brocher Summer Seminar (Geneva, Switzerland, 2022) and the 38th Annual Q Conference (Orlando, Florida; September 2022), where this work was presented. Finally, we extend our gratitude to the two anonymous reviewers, who provided constructive comments throughout the review process.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2023.116046.

References

Abelson, J., Lomas, J., Eyles, J., Birch, S., Veenstra, G., 1995. Does the community want devolved authority? Results of deliberative polling in Ontario. CMAJ (Can. Med. Assoc. J.): Canad. Med. Assoc. J. Journal de l'Association Medicale Canadienne 153 (4), 403–412.

Baker, R., Wildman, J., Mason, H., Donaldson, C., 2014. Q-ing for health—a new approach to eliciting the public's views on health care resource allocation. Health Econ. 23 (3), 283–297. https://doi.org/10.1002/hec.2914.

Baker, Rachel, McHugh, Neil, Mason, Helen, 2017. Constructing statement sets for use in Q methodology studies. In: Coast, J. (Ed.), Qualitative Methods for Health Economics. Rowman & Littlefield.

Banasick, S., 2019. KADE: a desktop application for Q methodology. J. Open Source Softw. 4 (36), 1360. https://doi.org/10.21105/joss.01360.

Barbour, A., 1997. Caring for Patients: A Critique of the Medical Model. Stanford University Press.

Barbu, C.M., 2014. Zoom: A Spatial Data Visualization Tool (2.0.6).

Barra, M., Broqvist, M., Gustavsson, E., Henriksson, M., Juth, N., Sandman, L., Solberg, C. T., 2019. Severity as a priority setting criterion: setting a challenging research agenda. Health Care Anal. 1–20. https://doi.org/10.1007/s10728-019-00371-z.

Brown, S.R., 1993. A Primer on Q Methodology, vol. 26.

Canadian Agency for Drugs and Technologies in Health, 2017. Guidelines for the economic evaluation of health Technologies: Canada (4th edition). https://www.cadth.ca/sites/default/files/pdf/guidelines_for_the_economic_evaluation_of_health_technologies_canada_4th_ed.pdf.

Cuppen, E., Breukers, S., Hisschemöller, M., Bergsma, E., 2010. Q methodology to select participants for a stakeholder dialogue on energy options from biomass in The Netherlands. Ecol. Econ. 69 (3), 579–591. https://doi.org/10.1016/j.ecolecon.2009.09.005.

Damar, E.A., Sali, P., 2022. Q Methodology: A Concise Overview [Chapter]. Methodological Innovations In Research And Academic Writing; IGI Global. https://doi.org/10.4018/978-1-7998-8283-1.ch001.

Daniels, N., 1985. Just Health Care. Cambridge University Press.

Diederich, A., Swait, J., Wirsik, N., 2012. Citizen participation in patient prioritization policy decisions: an empirical and experimental study on patients' characteristics. PLoS One 7 (5), e36824. https://doi.org/10.1371/journal.pone.0036824.

Dolan, P., 1998. The measurement of individual utility and social welfare. J. Health Econ. 17 (1), 39–52. https://doi.org/10.1016/S0167-6296(97)00022-2.

Dolan, P., Cookson, R., 2000. A qualitative study of the extent to which health gain matters when choosing between groups of patients. Health Pol. 51 (1), 19–30. https://doi.org/10.1016/s0168-8510(99)00079-2.

Dolan, P., Shaw, R., 2003. A note on the relative importance that people attach to different factors when setting priorities in health care. Health Expect. 6 (1), 53–59. https://doi.org/10.1046/j.1369-6513.2003.00210.x.

Drummond, M., Brixner, D., Gold, M., Kind, P., McGuire, A., Nord, E., 2009. Toward a consensus on the QALY. Value Health 12, S31–S35. https://doi.org/10.1111/j.1524-4733.2009.00522.x.

Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, G.L., Torrance, G.W., 2015. Methods for the Economic Evaluation of Health Care Programmes. Oxford University Press.

Engel, G.L., 1992. The need for a new medical model: a challenge for biomedicine. Fam. Syst. Med. 10, 317–331. https://doi.org/10.1037/h0089260.

Flick, U., 2013. The SAGE Handbook Of Qualitative Data Analysis. SAGE.

Green, C., 2009. Investigating public preferences on 'severity of health' as a relevant condition for setting healthcare priorities. Soc. Sci. Med. 68 (12), 2247–2255. https://doi.org/10.1016/j.socscimed.2009.03.020.

Gu, Y., Lancsar, E., Ghijben, P., Butler, J.R., Donaldson, C., 2015. Attributes and weights in health care priority setting: a systematic review of what counts and to what extent. Soc. Sci. Med. 146, 41–52. https://doi.org/10.1016/j.socscimed.2015.10.005.

Guest, G., M MacQueen, K., E Namey, E., 2012. Applied thematic analysis. SAGE publications, inc. https://doi.org/10.4135/9781483384436.

Gyrd-Hansen, D., 2004. Investigating the social value of health changes. J. Health Econ. 23 (6), 1101–1116. https://doi.org/10.1016/j.jhealeco.2004.02.002.

Hausman, D., 2019. The significance of 'severity'. J. Med. Ethics 45 (8), 545–551. https://doi.org/10.1136/medethics-2018-105058.

Hess, R.D., Hink, D.L., 1959. A comparison of forced vs. Free Q-sort procedure. J. Educ. Res. 53 (3), 83–90.

Hirose, I., Bognar, G., 2014. The Ethics of Health Care Rationing: an Introduction. Routledge.

Hofmann, B., 2002. On the triad disease, illness and sickness. J. Med. Philos.: A For. Bioethics Phil. Med. 27 (6), 651–673. https://doi.org/10.1076/jmep.27.6.651.13793.

Lakdawalla, D.N., Phelps, C.E., 2020. Health technology assessment with risk aversion in health. J. Health Econ. 72 (102346) https://doi.org/10.1016/j.jhealeco.2020.102346.

Linley, W.G., Hughes, D.A., 2013. Societal views on nice, cancer Drugs fund and value-based pricing criteria for prioritising medicines: a cross-sectional survey of 4118 adults in great britain. Health Econ. 22 (8), 948–964. https://doi.org/10.1002/hec.2872.

Magnussen, J., Aaserud, M., Granaas, T., Magelssen, M., Syse, A., Celius, E.G., Klovning, A., Syversen, I.D., 2015. På ramme alvor—alvorlighet og prioritering (p. 71) [Rapport fra arbeidsgruppe nedsatt av HOD]. Department of Health.

Malterud, K., 2019. Qualitative Meta synthesis: A Research Method for Medicine and Health Sciences, first ed. Routledge.

McHugh, N., Baker, R., Biosca, O., Ibrahim, F., Donaldson, C., 2019. Who knows best? A Q methodology study to explore perspectives of professional stakeholders and community participants on health in low-income communities. BMC Health Serv. Res. 19 (1), 35. https://doi.org/10.1186/s12913-019-3884-9.

McKeown, B., Thomas, D., 2013. Q methodology. SAGE Publications, Inc. https://doi.org/10.4135/9781483384412.

Millum, J., 2023. Should health research funding be proportional to the burden of disease? Polit. Philos. Econ. 22 (1), 76–99. https://doi.org/10.1177/1470594X221138729.

Mullen, P.M., 1999. Public involvement in health care priority setting: an overview of methods for eliciting values. Health Expect. 2 (4), 222–234. https://doi.org/10.1046/j.1369-6513.1999.00062.x.

National Institute for Health and Care Excellence, 2022. NICE Health Technology Evaluations: the Manual, p. 181.

NICE Citizens Council., 2008. Quality adjusted life years (QALYs) and the severity of illness. National Institute For Health And Care Excellence (NICE). http://www.ncbi.nlm.nih.gov/books/NBK401693/.

Nord, E., 1999. Cost-Value Analysis in Health Care: Making Sense Out of QALYS. Cambridge University Press.

Nord, E., Johansen, R., 2014. 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 Pol. 116 (2), 281–288. https://doi.org/10.1016/j.healthpol.2014.02.009.

Norheim, O.F., 2010. Priority to the young or to those with least lifetime health? Am. J. Bioeth. 10 (4), 60–61. https://doi.org/10.1080/15265161003697305.

Norheim, O.F., Allgott, B., Aschim, B., Førde, R., Gjul, G.K., Gundersen, T., Kakad, M., Kjellevold, A., Kvinsland, S., Melberg, H.O., Moen, A., Mæland, Ø., Olsen, J.A., Sjøli, S., 2014. Åpent og rettferdig—Prioriteringer i helsetjenesten (NOU 2014:12; p. 219). Helse- og Omsorgsdepartementet. https://www.regjeringen.no/no/dokumenter/NOU-2014-12/id2076730/.

Norheim, O.F., Emanuel, E.J., Millum, J., 2019. Global Health Priority-Setting: beyond Cost-Effectiveness. Oxford University Press. https://doi.org/10.1093/oso/9780190912765.001.0001.

Oddsson, K., 2003. Assessing attitude towards prioritizing in healthcare in Iceland. Health Pol. 66 (2), 135–146. https://doi.org/10.1016/s0168-8510(02)00211-7.

Olsen, J.A., 2013. Hva menes med «sykdommens alvorlighetsgrad»? Tidsskrift for Den norske legeforening. https://doi.org/10.4045/tidsskr.12.0659.

Rawls, J., 1999. A Theory Of Justice (Revised). Harvard University Press.

Reckers-Droog, V.T., van Exel, N.J.A., Brouwer, W.B.F., 2018. Looking back and moving forward: on the application of proportional shortfall in healthcare priority setting in The Netherlands. Health Pol. 122 (6), 621–629. https://doi.org/10.1016/j.healthpol.2018.04.001.

Riksdagsförvaltningen, 2018. Prioriteringar inom hälso- och sjukvården Kommittédirektiv 1992:8—Riksdagen. https://www.riksdagen.se/sv/dokument-lagar/dokument/kommittedirektiv/prioriteringar-inom-halso-och-sjukvarden_GGB18.

Rutgers, M.R., 2015. As good as it gets? On the meaning of public value in the study of policy and management. Am. Rev. Publ. Adm. 45 (1), 29–45. https://doi.org/10.1177/0275074014525833.

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., Jinks, C., 2018. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual. Quantity 52 (4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8.

Sayed, M.A., 2003. Psychotherapy of arab patients in the west: uniqueness, empathy, and “otherness”. Am. J. Psychother. 57 (4), 445–459. https://doi.org/10.1176/appi.psychotherapy.2003.57.4.445.

Schurer, M., Matthijsse, S.M., Vossen, C.Y., van Keep, M., Horscroft, J., Chapman, A.-M., Akehurst, R.L., 2022. Varying willingness to pay based on severity of illness: impact on health technology assessment outcomes of inpatient and outpatient drug therapies in The Netherlands. Value Health 25 (1), 91–103. https://doi.org/10.1016/j.jval.2021.08.003.

Shah, K.K., 2009. Severity of illness and priority setting in healthcare: a review of the literature. Health Pol. 93 (2/3), 77–84.

Skedgel, C., Henderson, N., Towse, A., Mott, D., Green, C., 2022. Considering severity in health technology assessment: can we do better? Value Health 25 (8), 1399–1403. https://doi.org/10.1016/j.jval.2022.02.004.

Solberg, C.T., Barra, M., Sandman, L., Hoffmann, B., 2023. Severity as a moral qualifier of malady. BMC Med. Eth. 24 (1) https://doi.org/10.1186/s12910-023-00903-2 article 1.

Stein, R.K., Perrin, E., Pless, I.B., Gortmaker, S., Perrin, J., Walker, D., Weitzman, M., In, R.C.O.C.I., 1987. Severity of illness: concepts and measurements. Lancet 330 (8574), 1506–1509.

Stephenson, W., 1935. Technique of factor analysis. Nature 136 (3434), 3434. https://doi.org/10.1038/136297b0.

Tenbensel, T., 2010. Virtual Special Issue Introduction: public Participation in health policy in high income countries - a review of why, who, what, which and where? https://doi.org/10.1016/j.socscimed.2010.08.005.

Tsuei, J.J., 1978. Eastern and western approaches to medicine. West. J. Med. 128 (6), 551–557.

Twaddle, A., 1994. Disease, illness and sickness revisited. In: Twaddle, A., Nordenfelt, L. (Eds.), Disease, Illness and Sickness: Three Central Concepts in the Theory of Health, pp. 1–18.

Ubel, P.A., 1999. How stable are people's preferences for giving priority to severely ill patients? Soc. Sci. Med. 49 (7), 895–903. https://doi.org/10.1016/S0277-9536(99)00174-4.

van der Weele, S., 2021. Thick concepts in social research: what, why, and how? Int. J. Qual. Methods 20 (16094069211066164). https://doi.org/10.1177/16094069211066165.

van Exel, J., Baker, R., Mason, H., Donaldson, C., Brouwer, W., 2015. Public views on principles for health care priority setting: findings of a European cross-country study using Q methodology. Soc. Sci. Med. 126, 128–137. https://doi.org/10.1016/j.socscimed.2014.12.023.

van Exel, N., de Graaf, G., 2005. Q Methodology—A Sneak Preview, p. 30.

Väyrynen, P., 2021. Thick ethical concepts. In: Zalta, E.N. (Ed.), The Stanford Encyclopedia of Philosophy (Spring 2021). Metaphysics Research Lab. Stanford University. https://plato.stanford.edu/archives/spr2021/entrieshick-ethical-concepts/.

Watts, S., Stenner, P., 2022. Doing Q Methodological Research. SAGE Publications Ltd. https://uk-sagepub-com.ezproxy.uio.no/en-gb/eur/book/doing-q-methodological-research.

Weinstein, M.C., Torrance, G., McGuire, A., 2009. QALYs: the basics. Value Health 12, S5–S9. https://doi.org/10.1111/j.1524-4733.2009.00515.x.

Wittenberg, E., Prosser, L.A., 2013. Disutility of illness for caregivers and families: a systematic review of the literature. Pharmacoeconomics 31 (6), 489–500. https://doi.org/10.1007/s40273-013-0040-y.

III

Charting public views on the meaning of illness severity

Mille Sofie Stenmarck[1, 2*]
David GT Whitehurst[3]
Rachel Baker[4]
Mathias Barra[1]

Affiliations
[1] The Health Services Research Unit – HØKH, Akershus University Hospital HF, Lørenskog, Norway
[2] Faculty of Medicine, University of Oslo, Norway
[3] Faculty of Health Sciences, Simon Fraser University, Canada
[4] Yunus Centre for Social Business and Health, Glasgow Caledonian University, Scotland
[*] Corresponding author

Abstract

Background: Illness severity is a central principle in multiple priority-setting frameworks, yet there is a paucity of research on public views regarding the meaning of illness severity. This study builds on the findings of a Q methodology study with members of the public that identified four general viewpoints on the meaning of illness severity. Here, we investigate the support for those viewpoints among the Norwegian population.
Methods: Following piloting, the online survey was distributed to a representative sample of the population (March to April 2023). The viewpoints from the earlier Q study were converted into vignettes: Lifespan, Subjective, Objective, and Functioning and Quality of Life (FQoL). The main task in the survey comprised ranking the vignettes and scoring them on a 0–10 visual analogue scale. We describe vignette alignment (from weak to strong) based on four categorisations (C1 to C4). C1 placed all respondents on their top scored vignette(s); C2 required a score of ≥7; C3 was designed to resolve ties; and C4 (which describes vignette *membership*) required a score of ≥7, a gap of two between vignettes scored ≥7, and did not allow ties.
Results: The survey was completed by 1174 individuals; those who completed in ≤3.5 minutes were excluded. Of the final sample (n=1094), 98.1% scored at least one vignette ≥7. In C1, 40.2% were aligned with Lifespan, 32.4% with FQoL, 28.9% with Objective, and 16.3% with Subjective. Using the C4 criteria, 55.4% did not have vignette membership, 13.6% had membership with Lifespan, 13.1% with Objective, 11.4% with FQoL, and 6.5% with Subjective.
Conclusions: There are several approaches to categorising vignette alignment, depending on the strength of alignment one wishes to describe. Across all analyses, Lifespan was scored highest and Subjective lowest. Our findings show that members of the public struggle to describe severity as one thing.

Keywords: healthcare, priority setting, severity, population survey, public views, Q2S

Introduction

To ensure transparent and systematic priority setting, healthcare systems depend on rigorous priority-setting frameworks. A key component of most priority setting frameworks is some form of health economic evaluation, usually based on costs and effects (Drummond et al., 2015). Health economic evaluation ensures that the costs of new treatments are weighed against their expected effect. Many countries also incorporate priority-setting principles related explicitly to moral value judgements, seeking to also ensure a fair and equitable distribution of health.

'Severity' is one such principle and is applied as an ethical decision-modifier in multiple healthcare jurisdictions, including the United Kingdom (UK) (National Institute for Health and Care Excellence, 2022), the Netherlands (Schurer et al., 2022), Sweden (Riksdagsförvaltningen, 1997), and Norway (Magnussen et al., 2015). As a priority-setting criterion, severity entails prioritisation of treatment options that may not meet standard cost-effectiveness thresholds by giving special consideration to conditions considered to be particularly severe. The severity criterion thus modifies standard cost-effectiveness analyses (Norheim, 2010), in line with egalitarian, prioritarian, or other non-consequentialist ethics (Barra et al., 2019; Ottersen et al., 2014).

Despite the implementation of severity as a priority-setting criterion in multiple healthcare systems, severity remains a contested concept and there is a lack of consensus on its meaning in the academic literature (Barra et al., 2019; Broqvist, 2018; Hausman, 2019; Nord & Johansen, 2014; Olsen, 2013). Furthermore, there are different policy operationalisations of severity at play across different jurisdictions. In Norway, the Netherlands, and the UK, severity is determined quantitatively based on the estimated loss of quality-adjusted life years (QALYs), with different QALY-based estimations in use in the three countries (Ministry of Health and Care Services, 2020; National Institute for Health and Care Excellence, 2022; Reckers-Droog et al., 2018). In Sweden, severity is taken into account qualitatively, using descriptive severity levels (Riksdagsförvaltningen, 2018).

A participatory turn in health-related research and policy in recent years has led to increased emphasis on public views in policy-making processes, with decision-making bodies increasingly expected to include general public representatives in decision-making processes (Abelson et al., 2013; Baker et al., 2021). There are multiple population-level studies eliciting public views on the relevance of severity as a priority-setting criterion. These demonstrate that there is broad support for severity as a relevant concern in priority setting (Gu et al., 2015; Shah, 2009; Skedgel et al., 2022). However, while these studies highlight the degree to which members of the public consider severity to be a relevant concern, they do not explore what severity is taken to *mean*. In a recent study, Stenmarck and colleagues explored public views on the meaning of illness severity by conducting group interviews with members of the Norwegian general public (Stenmarck et al., 2023). Using Q methodology to generate distinct and shared viewpoints on severity, four different viewpoints on severity were identified. These viewpoints demonstrate that severity is a complex concept for members of the public, with contrasting views on its meaning.

Q methodology combines qualitative and quantitative techniques to render rich descriptions of subjective views on a given topic. Being in-depth and exploratory, Q studies provide detailed information about the views of participants (Brown, 1993; Watts & Stenner, 2022). However, Q studies are typically conducted with relatively small groups of respondents and do not provide information about the distribution of viewpoints across a population (Brown, 2002). The recent study by Stenmarck and colleagues provides knowledge about different views on illness severity—but if policy makers are to make use of information about public views on severity, they also need to know the extent to which those views are shared by the population and whether different sub-populations think differently about these views. Such knowledge may support healthcare policies referring to severity that align with the views of the population. This study addresses the knowledge gap on how different accounts of illness severity are distributed among the Norwegian population.

Methods

Design

This study is part of the 'Severity in priority setting' (SEVPRI; Norwegian Research Council no. 303724) project and builds on results from the aforementioned study by Stenmarck and colleagues (Stenmarck et al., 2023).

Four viewpoints on illness severity were identified and described in the original Q study, labelled (i) 'natural lifespan', (ii) 'severity is subjective', (iii) 'objective measures and triage', and (iv) 'functioning and quality of life' (Stenmarck et al., 2023). There are several ways of designing a survey with Q2S methods (i.e., methods to go from a Q study to a survey) (Baker et al., 2010). The survey presented in this study uses summary descriptions (van Exel et al., 2011), where four vignettes were developed to represent the viewpoints identified by Stenmarck and colleagues. The vignettes included in the survey were developed to contain details of the most central features of the four viewpoints; effectively, each vignette is a brief summary of one of the viewpoints. Throughout this paper, the four vignettes are labelled, Lifespan, Subjective, Objective, and Functioning and Quality of Life (FQoL). As the vignettes needed to be suitable for a survey (i.e., not too lengthy), each vignette is approximately 100 words in length. The vignettes used in the survey are presented in Table 1. Further information about vignette development is provided in Supplementary Material A.

Preliminary drafts of the vignettes were discussed with SEVPRI's Advisory Board for feedback on wording and intelligibility. A version of the full survey was then piloted with members of a user panel at Akershus University Hospital (consisting of members of the public) and colleagues at Akershus University Hospital, to receive feedback on how well people understood the tasks and the accompanying explanations.

The online survey was administered by Norstat (www.norstat.no), a market research company. Respondents were informed that the survey should take approximately 15 minutes to complete and were encouraged to complete it in an undisturbed and quiet environment. The survey comprised six sections: (i) an introduction, (ii) a statement-ranking 'triplets' task (which is not part of the analysis reported here), (iii) presentation of the vignettes, (iv) a vignette ranking task, (v) a vignette scoring task, and (vi) a series of post-survey questions. Figure 1 illustrates the design. The full survey (in Norwegian) can be accessed here: https://web.norstatsurveys.com/survey/selfserve/53c/2308137.

The survey's introduction briefly outlined the context of severity in priority-setting frameworks. Following the triplets task, each vignette was presented, in random order, one after another. The subsequent two-step vignette ranking and scoring task was the main section of the survey. For the ranking task, the vignettes were presented in the same order as they were presented individually, but this time shown together on screen, and respondents were asked to rank them from "most like my view" to "most unlike my view" (top to bottom.) Participants did this by dragging and dropping the vignettes into vertically placed boxes. For the scoring task, the vignettes were presented again, this time in the order that the respondents had previously ranked the vignettes. Participants were asked to place the vignettes on a vertical visual analogue scale (VAS), anchored at the top by "completely descriptive of my view" and at the bottom by "completely different from my view". The VAS had nine evenly spaced non-numbered tick marks (with smaller evenly spaced tick marks between each of these). In the data set provided by Norstat, these were converted into eleven uniformly spaced intervals, coded from zero to ten.

Demographic information about respondents was captured by the information Norstat has about panel members. Respondents were presented with their recorded demographics already filled in during the introduction of the survey and had the option to revise their information before continuing. In the final section, respondents were offered the opportunity to provide their own description of severity, either based on the provided vignettes or by drafting a new description (these descriptions are not part of the analysis reported here). They were then asked to provide feedback on the survey, both in terms of the difficulty of performing the tasks and their opinion on the quality of the instructions. Details on the post-survey questions and responses to these are presented in Supplementary Material B.

Respondents

Respondents were sampled from Norstat's survey panel of members of the Norwegian general population aged 18 to 99 years. Norstat delivers target quotas for age (represented by five age categories: 18–29, 30–39, 40–49, 50–59, and 60–99), sex ('male' or 'female'; in Norway only two sex categories are recognised in law), and region (dividing Norway into the regions of East, South, West, Mid, North, and the capital, Oslo), set to be nationally representative according to data from Statistics Norway. Norstat also provides postal codes for home addresses, which we used to compute the Statistics Norway centrality index. The centrality index grades rural-to-urban municipality from 0 to 1000, deriving six centrality classes (Høydahl, 2020). Gross household income was also provided. We used open source data from Statistics Norway to explore how representative our sample was compared with the Norwegian general population (Høydahl, 2020).

The target sample size was 1,000 respondents, in addition to a pilot of 100. Adaptations to the survey were made following the pilot phase; these changes were substantial enough to merit a second pilot, where an additional 100 responses were collected. All respondents in the final survey were entered into a lottery for two cash prizes (gift cards) of NOK 5,000 (~€450).

Analysis

Our analysis treats the vignette ranking and vignette scoring sections of the survey as a single, two-step task, where the VAS scores define respondents' alignment with the vignettes. These continuous data tell us something about the degree of alignment each respondent has with each of the vignettes. In addition to describing the degree of alignment for each respondent across all four vignettes, we wished to explore how respondents could be linked to the vignettes. Placing respondents into categories, according to their alignment, affords opportunities to describe the prevalence and distribution of views. This requires rules/criteria to define category assignment. Any selected criteria are somewhat arbitrary and so we present data according to different categorisation rules, which are intended to represent different strengths of alignment, from relatively weak alignment to relatively strong alignment.

Four rule-based categorisations were defined to explore the strength of respondents' alignment with the vignettes. These categorisations, described below, are illustrated in Figure 2. Based on a qualitative interpretation of the VAS, a trisection of the 0–10 range of the VAS was implemented, with 0–3 reflecting a respondent's indication that the respective vignette does not align with their views (herein, 'disapproval'), scores from 4-6 reflecting a neutral position ('neutral'), and scores from 7–10 reflecting alignment with the respondent's views ('alignment').

The first categorisation (C1) aligns respondents with the vignette(s) they scored highest (i.e., ties are possible, should a respondent assign their highest VAS score to more than one vignette). A second categorisation (C2) represents stronger vignette alignment, placing respondents on a vignette if there is a VAS score ≥ 7 (as with C1, ties are permitted). As the top and bottom of the VAS read "completely descriptive of my view" and "completely different from my view", C2 regards respondents who place their top vignette(s) below 7 as not having alignment with any of the vignettes, i.e., compared with C1, C2 requires a more explicit statement of alignment to a vignette.

The third categorisation (C3) provides an alternate 'stronger than C1' category. C3 does not require a respondent to score above a certain point on the VAS, but ties are not allowed. All respondents are considered aligned with one vignette only; if a respondent tied their top vignettes in the scoring task, the tie is resolved by reverting to the respondent's vignette ranking. The fourth categorisation (C4) is the strictest vignette alignment rule, which we refer to as defining vignette 'membership'. Vignette alignment under C4 requires a VAS score of ≥7 and excludes ties. If more than one vignette was given a VAS score ≥7, a gap of at least two points on the VAS (i.e., 7 and 9, 7 and 10, or 8 and 10) is required between the vignettes. This requirement has been applied in other Q2S studies and allows for a more confident assertion of vignette membership (Mason et al., 2018). While C3 aligns all respondents with one vignette, C4 permits non-alignment; both place each respondent in one exclusive category. We computed χ^2-test statistics for the univariate distributions of C3- and C4-defined alignement, to explore the distribution of respondents across the vignettes in these two categorisations.

We fitted four simple linear regression models, one for each vignette, with the dependent variable being the respondents' VAS score and independent variables being the respondent characteristics (age, sex, gross household income, education level, etc.) to identify possible relationships between vignette alignment and sociodemographic characteristics. These analyses were exploratory; no hypotheses were defined prior to fitting the regression models, and we therefore do not specify any significance level. As such, these analyses are exploratory, and should be interpreted with care. We report standard regression output, including p-values, but refrain from discussing the results as significant or non-significant, instead providing a descriptive summary of the results.

Respondents' engagement with the survey was explored prior to finalising the analytic sample. This was done by analysing completion time, with the intention of excluding participant who completed the survey 'too quicky'. Here, 'completion time' refers to the time spent on the triplets task and the ranking and scoring tasks only. The completion-time cut-off was determined by assessing respondents' engagement with the survey through non-interaction or likely random response patterns. For example, respondents who did not rearrange any of the statements during the triplets task nor any of the vignettes during ranking were unlikely to be actively engaging with the survey. Another possible proxy for non-engagement could be inconsistent ordering of the vignettes in the ranking and scoring tasks. It is important to note that at the individual level, 'inconsistency'—defined as scoring the vignettes on the VAS task in a manner not consistent with their previous ranking—does not imply that a respondent did not understand the task or were disengaged with the survey. There was no 'go back' option available once a respondent arrived at the scoring task, and an inconsistency thus defined could result both from making a mistake during ranking, or possibly following thoughtful reconsideration during the scoring task. However, at the aggregate level, an analysis of the association between response times and both the degree of inconsistency (randomness between ranking and scoring) and the likelihood of non-interaction was used to inform a completion-time cut-off for exclusion from further analyses.

Finally, we explored respondents' interaction with the survey by analysing how they rearranged the randomly ordered vignettes in the ranking task, by respondents' use of the VAS, and by performing pairwise correlation analysis of the vignettes. Further details on these analyses are reported in Supplementary Material C.

All data analyses were conducted with R Studio (2023-06-1 Build 524) running on R version 4.3.0 (2023-04-21) (RStudio Team, 2020)

Ethics

The Regional Committee for Medical and Health Research Ethics (Regional Ethics Committee South-East B) advised that the aims and objectives of SEVPRI were not regulated by the Health Research Act (Helseforskningsloven, 2008). The project's protocol was evaluated by the Data Privacy Officer at Akershus University Hospital, who advised that the study could be conducted (PVO nos. 20_200 and

21_200). Akershus University Hospital and the Principal Investigator (MB) are responsible for project oversight. Norstat collected consent from all respondents via their panel standards, and respondent data was handled according to Norstat confidentiality and General Data Protection Regulation (GDPR) requirements (Norstat, 2023).

Results

The two online pilot phases were performed between December 2022 and January 2023, with 99 respondents in the first pilot and 134 in the second. Analysis of these pilot data demonstrated a high rate of inconsistencies (~40%) between the ranking and the scoring tasks. This was attributed to the scoring task being difficult to perform within the online program, with the vignettes difficult to drag and place on the VAS. Improvements were made and a notification was added. For respondents who altered the order of the vignettes from the ranking task when completing the scoring task, the notification made them aware of this and informed them that they could re-drag the vignettes to the scale if this was done by mistake. Following these changes, the rate of inconsistencies dropped to ~20%.

A total of 1,174 respondents completed the final survey. Relatively high frequencies of inconsistencies were observed for respondents who completed the survey in a few minutes, with a notable drop off (i.e., more consistent responses) after 3.5 minutes. Two authors also completed the survey and were unable to do this in 3.5 minutes, despite their familiarity with it. Accordingly, those who completed the survey in less than 3.5 minutes (n=80) were excluded; a final sample of 1094 responses were retained for further analyses. Of the 1,094 responses, 84.9% provided consistent orderings.

Respondent characteristics are reported in Table 2, along with comparable statistics for the Norwegian population. The study sample was similar to the Norwegian population regarding age, region, and centrality, but overrepresentation of women and individuals with higher levels of education. Information regarding gross household income (HHI) was missing for 174 respondents (122 responded ‘prefer not to say’ and 52 responded ‘don’t know’). There was a correlation between missing HHI and sex, age, and education. We therefore imputed the median HHI of respondents matched for sex, age, and education for the 174 respondents without information on HHI.

The four regression models used to explore associations between respondent characteristics and vignette alignment are reported in Table 3. The included independent variables explain little of the observed variation in respondents’ scoring of the four vignettes (adjusted R^2 in the range 0.014–0.034). However, we found some evidence for associations between respondent characteristics and VAS scores for the four vignettes: women appear more likely to score Lifespan and FQoL higher; high HHI appears to be associated with higher VAS scores for Lifespan and Objective; and there is an association between higher scores for Subjective and low HHI, low education, and higher age.

In response to the post-task questions, respondents reported that, overall, the ranking and scoring task was difficult to perform, but the survey was well explained (see Supplementary Material B).

Vignette alignment

Prior to presenting results by categorisation—which is the primary focus of the study—general observations from the study sample were that (i) 84.7% of respondents awarded their highest VAS score to one vignette, (ii) 28.9% scored at least one vignette at the maximal score (10) indicating complete alignment with those vignettes, (iii) 73.1% gave at least one vignette a VAS score of 9 or 10, and (iv) 98.1% scored at least one vignette within the alignment range (a VAS score between 7 and 10). Only 1.0% of respondents gave their least preferred vignette a VAS score of zero. Further details regarding the respondents' use of the VAS scoring range and descriptive statistics and correlation analysis for the VAS scores are presented in Supplementary Material C.

Categorisation 1 (C1)

The setup of the scoring task allowed respondents to award different vignettes the same VAS score, creating the possibility of ties. In such a case, both vignettes could represent the respondent's 'most endorsed' vignette (or 'least disliked' vignette), and the respondent would likely be indifferent between those vignettes and accepting a random selection between the tied vignettes. The Venn diagram presented in Figure 3 illustrates the distribution of respondents across the vignettes, where the 'outskirts' of the Venn diagrams represent the percentage of respondents who gave the respective vignette their single highest VAS score, and the areas with overlap represent the percentage of respondents who had different combinations of tied vignettes.

Lifespan was scored highest by 440 respondents (40.2%), FQoL by 354 (32.4%), Objective by 316 (28.9%), and Subjective by 178 (16.3%). These numbers sum to more than 100% of the study sample size due to the acceptance of ties within C1. A total of 927 (84.7%) respondents gave a single vignette their highest VAS score (i.e., the non-overlapping sectors of Figure 3), most frequently the Lifespan vignette (n=316 (28.8%)). Respondents tied two vignettes on 142 (13.0%) occasions. The most frequent two-vignette ties were Lifespan and FQoL (n=50; 4.6%), and Lifespan and Objective (n=31; 2.8%). The least frequent two-vignette tie was Subjective and Objective (n=3; 0.3%). These two vignettes also distinguished themselves as being the least tied vignettes when accounting for ties of three or four. Eleven (1.0%) respondents tied at least Subjective and Objective. Twenty-five (2.3%) respondents tied three or four vignettes.

Categorisation 2 (C2)

In C2, respondents could be assigned between 0 (where no vignette was given a VAS scored ≥7) and 4 (where all vignettes were given VAS scores ≥7) vignettes. As illustrated in Figure 4, 914 (83.6%) respondents gave a single vignette their highest VAS score. As in C1, Lifespan was the highest scored single vignette (28.5%), followed by FQoL (22.1%) and Objective (22.0%), and finally Subjective (10.9%). A difference between the classifications of C1 and C2 is that 21 (1.9%) respondents are not assigned to a vignette in C2 (because 21 respondents did not give any vignette a score ≥7). Most of the unassigned respondents under C2 (13 of 21) were in one of the single-vignette sectors of C1.

Categorisations 3 and 4 (C3 and C4)

C3 resembles C1 in that respondents are considered aligned with their top-scored vignette, irrespective of where the vignette was placed on the VAS. However, if they tied their top vignettes, C3 uses rankings to break the ties. C4 places additional criteria to that operationalised by C3, as described in the methods section (and Figure 1). Venn diagrams similar to Figure 3 and Figure 4 are not relevant for the analysis of C3 and C4 because there are no ties. Instead, results using the C3 and C4 classifications are presented in Table 4, disaggregated by respondent characteristics. Results for the full sample demonstrate that, by C3 criteria, the highest proportion of respondents (34.7%) are aligned with Lifespan, followed by FQoL (27.0%), then Objective (25.0%), and finally Subjective (13.3%). By C4 criteria, 55.4% are not assigned to a vignette. C4 has the strictest criteria—what we consider to describe vignette *membership*. Of those aligned with vignettes, 13.6% have membership with Lifespan, 13.1% with Objective, 11.4% with FQoL, and 6.5% with Subjective.

Looking at vignette alignment and respondent characteristics, men report stronger alignment with Subjective than women for both the C3 and C4 criteria. In C3, 17.4% of men are aligned with Subjective, compared with 9.6% of women; in C4, 9.4% of men have membership with Subjective, compared with 3.7% of women. There also appears to be a systematic relationship between alignment with Subjective and the characteristics of education level and HHI, with 18.2% of those with 'undergraduate-level and below' education in C3 (compared with 9.2% of those with graduate-level education), and 9.6% with undergraduate or below having membership with Subjective by C4 (compared with 4.1% of those with graduate-level education). Under the C3 criteria, alignment with Subjective was 17.5% for those in the lowest HHI category, compared with ~10–12% in the three higher income categories. Also under the C3 criteria, 45.6% of those living in the North report alignment with Lifespan, which is higher than the proportion for any other vignette (based on C3 in any other area of Norway). There does not appear to be any general association between age and vignette alignment. None of the listed characteristics appear to explain vignette alignment for respondents categorised by C4 as non-aligned.

Discussion

A participatory turn in health-related research and policy has made the views of the public increasingly important in priority-setting processes (Baker et al., 2021; Chalkidou, 2012). Because illness severity is a priority-setting principle in multiple healthcare systems, how severity is interpreted and operationalised in priority-setting frameworks has considerable impact on the distribution of healthcare resources. Knowledge of public views on severity is therefore relevant to policy makers. Four different viewpoints on the meaning of severity were identified in a previous study (Stenmarck et al., 2023). This paper has presented an analysis of the distribution of those viewpoints across the Norwegian population, converting the viewpoints into vignettes and using online survey methods.

Analysis of 1094 responses (following the exclusion of 80 respondents who completed the survey in under 3.5 minutes) led to several important findings. First, with 98.1% of respondents scoring at least one vignette at 7 or above, it seems the viewpoints on severity identified in the original Q study resonate with the wider population. Second, that most respondents scored more than one vignette at 7 or above demonstrates that severity, to members of the public, means several different things. Lifespan and FQoL were the two vignettes most tied, while Subjective and Objective are infrequently tied. These findings are intuitive and lend support to the face validity of the survey. Lifespan and FQoL both emphasise functioning in everyday life and the importance of being able to live a somewhat 'normal' life. Subjective and Objective represent directly opposing views on severity (Stenmarck et al., 2023). While the Subjective view centres on severity as something which is defined personally and determined by the individual, the Objective account sees severity as a measurable notion, best determined by health personnel or other professionals. Overall, the high VAS scores across the vignettes demonstrate that no single vignette can represent what members of the public consider illness severity to mean. This supports findings from several qualitative studies, where members of the public considered severity to be a complex concept that does not necessarily mean 'one thing' (Broqvist et al., 2018; Magnussen et al., 2015; Stenmarck et al., 2023).

Third, across all analyses, Lifespan garnered most support, followed by Objective and FQoL. According to the Lifespan view, severity is tied to how illness affects the natural course of life, and whether illness impedes one's ability to reach important milestones and to live the life one wishes to. Lifespan can be compared with the 'fair innings' argument in health economics, centred on the notion that everyone is entitled to a certain amount of quality-adjusted life expectancy (Williams, 1997). Support for the Lifespan view may reflect that members of the public consider illness severity (specifically) and health (generally) to be more than one's level of functioning, as FQoL is oriented around, or biological markers, as Objective centres on, and more as an expression of what kind of life one is able to lead. One could deduce from this that the fair innings argument resonates with a considerable portion of the wider population. Across all analyses, a minority of respondents were aligned with the Subjective view, which is distinguished by rejection of severity criteria—a notion which does not appear to resonate strongly with the wider population.

Finally, while we did not have any *ex ante* hypotheses about a relationship between sociodemographic qualities and the vignettes, our findings indicate that men, those with lower education, and those with lower HHI are more likely to be aligned with the Subjective view, compared with others. Subjective is thus tied to lower socioeconomic status, measured here by education and HHI. Socioeconomic status is an important concern in health policy due to the socioeconomic gradient in health. The socioeconomic gradient is a well-established phenomenon stating that individuals with lower socioeconomic status have increased disease burden and mortality compared to those with higher socioeconomic status (Bonaccio et al., 2020; Marmot, 2004). This means that individuals with lower socioeconomic status are more likely to have lived experience with illness. The Subjective vignette is centred on the notion that those with lived experience ought to determine what counts as severe illness. Our findings suggest that those likely to have lived experience with illness (i.e., those with lower socioeconomic status) are more likely to think that experience should count for something. The Subjective vignette is also the one that most starkly contrasts current policy operationalisations of severity in countries using severity as a priority-setting criterion, which generally determine severity according to more objective measures of illness based on loss of quality and quantity of life (Ministry of Health and Care Services, 2020; National Institute for Health and Care Excellence, 2022; Reckers-Droog et al., 2018; Riksdagsförvaltningen, 2018). This ties our findings to the ongoing debate on the role of lived experience in health policy, and to what degree lived experience should inform understandings of illness (Reynolds, 2022; Toombs, 1995).

Strengths and limitations

This study has three main strengths. First, to our knowledge, this is the largest exploration of public views on the meaning of severity. While this study was conducted in Norway, our findings are of value in an international context, given the relevance of severity as a decision-making criterion in multiple healthcare systems. Second, the survey was based on an in-depth analysis of views on severity, building on the qualitative and quantitative elements of a previous Q study (Stenmarck et al, 2023). Third, this paper reports one of few Q2S vignette studies (Baker et al., 2010), thus contributing to the small pool of Q2S studies and providing an example of alternative approaches to exploring public views on complex concepts. Finally, we demonstrated that respondents described the ranking and scoring task as difficult to perform, while also indicating the survey instructions were clear (see Supplementary Material B). It should be no surprise that respondents find complex tasks difficult to complete. This does not mean that multi-statement vignettes are unsuitable for surveys, rather it serves as confirmation of the importance of extensive piloting and clear, informative instructions at the outset of the survey (e.g., letting respondents know that the survey requires concentration and give them the opportunity to find a peaceful setting to complete it, allowing for the focus needed to provide considered responses).

As with any study, there are limitations. First, some of the underlying meaning in the viewpoints from the original Q study may have been lost in the summary vignettes. The process of converting the results of the original Q study into a survey also introduces risk of researcher bias (Guest et al., 2012). We made efforts to avoid this by striving to identify our biases throughout the course of the study and adopting a reflexive approach throughout the research process (Flick, 2013). Second, survey participants were incentivised by entry into a cash lottery. Such an incentive could introduce bias if people complete the requested tasks without sufficiently engaging with the task (Laguilles et al., 2011). Potential evidence of this was seen in two ways. Of the 1174 complete responses, 99 (8.4%) participants did not rearrange any of the vignettes in the ranking task, which is twice the expected 4.2% (see Supplementary Material C). After excluding the 80 (6.8%) participants that completed the main task in less than 3.5 minutes, the share of respondents that displayed this non-engaged ranking behaviour was reduced to 6.9% (76 of the 1094 retained respondents).

Third, prior to analysis and design of the C1–C4 categorisations, we made an informed judgment about the meaning of vignette placement on the VAS. However, while the VAS is often interpreted as a 'linear scale', there is a possibility of non-linearity in respondents' use of the VAS scale (Hartmannsgruber & Silverman, 2000; Myles et al., 1999). For example, it is possible that respondents felt that the difference between a score of a 9 and a 10 was in some sense more significant, or important, than the difference between a 5 and a 6. Finally, the study sample had an over-representation of women and individuals with higher education. It is possible that additional characteristics, such as political affiliation, religious beliefs, or health state, might have further explanatory strength. We did not have the necessary ethical approvals to ask respondents about such characteristics and were wary of survey fatigue (Backor et al., 2007).

Conclusion

This study reports the largest study so far investigating public views on the meaning of severity in the context of healthcare. The results of this study provide valuable information for decision makers in healthcare, particularly in healthcare systems incorporating (or considering incorporating) severity as a priority-setting criterion. Methodologically, we demonstrate that there are several approaches to categorising vignette alignment, depending on the strength of alignment one wishes to describe.

While Lifespan is the vignette that was scored highest across the different analytical approaches, we demonstrate that members of the public associate severity with several different, often contrasting, meanings. While our results show that none of the sociodemographic sub-groups we could isolate displayed clear associations to the vignettes (independent variables explained very little of the variation in VAS scores), there is some evidence of possible social gradients worthy of further research; in particular the question of whether individuals with poorer health—which is associated with lower income and lower education—align themselves more closely with Subjective than other groups.

This study shows that members of the public struggle to describe severity as 'one thing'. Almost all respondents expressed alignment with at least one vignette, and many aligned themselves with more than one. Moreover, very few endorsed all the vignettes, suggesting that there is substantial disagreement in the population about what characterises illness severity. We also argue that the meanings members of the public associate with severity are not necessarily reconcilable with current policy operationalisation of severity. Decision makers ought to bear this plurality of views in mind in future iterations of priority-setting policies involving the concept of severity.

Acknowledgements

The authors would like to thank all those who took part in the pilot stages and the final survey. We would also like to thank Professor Job van Exel and Professor Helen Mason for their valuable input on survey design.

Author contributions

Design of the study was a collaborative effort between all authors. The initial draft of the vignettes was written by MSS. Survey development was led by MSS and MB, with input from all authors. Analysis was conducted by MSS and MB. Interpretation of results was discussed between all authors. The first draft of the paper was written by MSS. All authors critically reviewed the paper for intellectual content. All authors have read and approved the paper for submission.

Funding

SEVPRI is funded by the Research Council of Norway (ref. no. 303724). All authors are fully or partially supported by this grant.

Conflicts of interest

None of the project collaborators have any conflicts of interest, nor any financial interests in the project.

Supplementary material

Supplementary Material A: Vignette development
Supplementary Material B: Survey feedback
Supplementary Material C: A summary of (i) respondents' rearranging of randomly ordered vignettes, (ii) the use of the VAS scoring range, and (iii) descriptive statistics and correlation analysis for the VAS scores

Tables

Table 1 The four vignettes used in the survey, which are summary descriptions of the four viewpoints identified in the original Q study (Stenmarck et al, 2023).[a]

Vignette I: Lifespan (93 words)
I think severity is about how health problems affect the natural course of life, and affect the natural development of life – and especially if it affects the young. Death isn't necessarily severe, especially when you're old, and there are things that are more severe than death. Pain isn't necessarily severe either. But things like mental illness and loss of dignity, that is severe. Severity is first and foremost when illness affects the natural course of life's different phases, and takes one's possibility to experience what you should be allowed to expect from life.
Vignette II: Subjective (112 words)
I think severity is almost entirely about how one experiences health problems, and severity can't be tied to any specific diagnoses or conditions. Severity depends completely on the experienced situation, and what one feels is severe for oneself and one's life. You can't define severity objectively. There is no right answer to what severity is, but when you are affected at a young age, or by something that just gets worse and worse, or leads to a loss of dignity, that makes it more severe. And it's maybe more severe if it affects a parent who is responsible for a child. But it's hard to say anything definitive about what severity is.
Vignette III: Objective (91 words)
I think severity has to be defined by some objective measures, like age, diagnosis, prognosis, and urgency. It can't be up to each individual to decide what is severe and what isn't. We need some criteria, and health personnel or other experts should be involved in making those criteria. Both mental illness and pain can be severe, or not being able to work, but severity is not about what someone believes, feels, or thinks. The degree of severity is decided by objective facts about the condition or diagnosis that one has.
Vignette IV: Functioning and quality of life (80 words)
I think severity depends on how it affects your day-to-day life: if it alters your functioning at work, at home, your ability to participate in society, enjoy your hobbies and things like that. Still, not just anything can be severe. There are clearly some objective links between a health problem (mental or physical) and how your quality of life is affected. But severity has to be measured by how it affects your quality of life and your levels of functioning.

[a] In the survey, the vignettes were presented in Norwegian (see Supplementary Material A). The English translations are provided for publication purposes only.

Table 2 Respondent characteristics with comparable statistics for the Norwegian general population. Values are numbers (percentages) unless stated otherwise.

Characteristic	**Final sample (n=1094)**	**Norwegian population (percentage only)**
Sex		
Male	530 (48.5)	50.2
Female	564 (51.5)	49.8
Age		
18-29	234 (21.4)	20.3
30-39	198 (18.1)	16.9
40-49	184 (16.8)	17.3
50-59	198 (18.1)	16.4
60-99	280 (25.6)	29.2
Region		
Oslo (HS)	163 (14.9)	13.0
East (HS)	323 (29.5)	29.8
South (HS)	130 (11.9)	13.6
West (HW)	227 (20.7)	20.5
Mid-Norway (HM)	161 (14.7)	13.8
North (HN)	90 (8.2)	9.3
Highest completed education[a]		
Elementary	49 (4.5)	23.6
Upper elementary	216 (19.7)	36.1
Apprenticeship	200 (18.3)	31.4
Undergraduate	301 (27.5)	25.1
Graduate	316 (28.9)	11.5
Other/none	12 (1.1)	0.6
Centrality class[b]		
Class 1 (most urban)	233 (21.3)	19.0
Class 2	312 (28.5)	25.3
Class 3	292 (26.7)	25.4
Class 4	150 (13.7)	16.5
Class 5	77 (7.0)	9.4
Class 6 (most rural)	30 (2.7)	4.5

HM, Health Mid; *HN*, Health North; *HS*, Health South; *HW*, Health West. Norway's health trusts are divided into four health regions: South-East, corresponding to the Norstat regions Oslo, East, and South, and the remaining three (Middle, West, and North) corresponding to the Norstat-defined regions.

[a] The proportions for the highest level of completed education were estimated from SN:09439 in Statistics Norway (Høydahl, 2020).

[b] The proportions for centrality classes were estimated from Table 5.2 in Statistics Norway (Høydahl, 2020), where a centrality index grades rural-to-urban municipality from 0 to 1000, deriving six centrality classes.

Table 3 Four linear regression models (one for each vignette), demonstrating relationship between vignette alignment and sociodemographic characteristics.

Characteristic	**Lifespan**	**Subjective**	**Objective**	**FQoL**
Age	-0.009*	0.016**	-0.016*	-0.007
	(0.004)	(0.005)	(0.005)	(0.004)
Male sex (ref=female)	-0.503***	0.231	0.004	-0.565***
	(0.144)	(0.166)	(0.177)	(0.140)
Region (ref= East)				
Oslo (HS)	-0.067	-0.528	0.596*	-0.383
	(0.244)	(0.282)	(0.301)	(0.238)
South (HS)	0.157	-0.054	0.300	-0.479*
	(0.247)	(0.285)	(0.304)	(0.240)
West (HW)	0.026	-0.352	0.251	-0.195
	(0.205)	(0.236)	(0.252)	(0.199)
Mid-Norway (HM)	0.050	0.005	0.460	-0.207
	(0.233)	(0.269)	(0.287)	(0.227)
North (HN)	0.346	-0.233	0.448	-0.522
	(0.303)	(0.349)	(0.373)	(0.295)
Education (ref=undergraduate and below)	-0.014	-0.784***	0.428*	0.092
	(0.151)	(0.174)	(0.186)	(0.147)
Centrality class[b]	-0.103	-0.028	-0.052	-0.016
	(0.071)	(0.082)	(0.088)	(0.069)
Log(HHI)[c]	0.338**	-0.390**	0.478**	0.029
	(0.129)	(0.148)	(0.159)	(0.125)
Constant	*5.734****	*7.362****	*3.025***	*7.461****
	(0.842)	*(0.972)*	*(1.038)*	*(0.821)*
Observations	1094	1094	1094	1094
R^2	0.023	0.043	0.029	0.026
Adjusted R^2	0.014	0.034	0.020	0.017
Residual standard error (df=1080)	2.318	2.674	2.856	2.258
F Statistic (df=11; 1080)	2.502**	4.891***	3.177***	2.895**

[a] *HM,* Health Mid; *HN*, Health North; *HS*, Health South; *HW*, Health West. Norway's health trusts are divided into four health regions: South-East, corresponding to the Norstat regions Oslo, East, and South, and the remaining three (Middle, West, and North) corresponding to the Norstat-defined regions.

[b] Centrality class (a centrality index grades rural-to-urban municipality from 0 to 1000, deriving six centrality classes) is used as a continuous variable in the regression.

[c] Log(HHI) = log-gross household income.

*p<0.05; **p<0.01; ***p<0.001

Table 4 Respondent characteristics matched with C3 and C4 criteria. Values are percentages.

Characteristic	Lifespan		Subjective		Objective		FQoL		Not assigned
	C3	C4	C3	C4	C3	C4	C3	C4	C4
All	34.7	13.6	13.3	6.5	25.0	13.1	27.0	11.4	55.4
Sex*,**									
Male	32.6	12.6	17.4	9.4	26.0	14.0	24.0	10.6	53.4
Female	36.7	14.5	9.6	3.7	23.9	12.2	29.8	12.2	57.3
Age ^{NS,NS}									
18-29	37.2	12.4	10.7	6.0	21.8	9.8	30.3	11.5	60.3
30-39	34.8	10.6	9.6	4.0	30.3	16.2	25.3	11.1	58.1
40-49	34.8	15.2	14.1	5.4	24.5	12.5	26.3	11.4	55.4
50-59	29.8	11.6	17.7	7.1	22.2	11.1	30.3	14.1	56.1
60-99	36.1	17.1	14.6	8.9	26.1	15.4	23.2	9.6	48.9
Region[a] ^{NS,NS}									
Oslo (HS)	36.2	13.5	10.4	6.1	29.4	14.7	23.9	9.8	55.8
East (HS)	31.9	12.7	15.2	7.4	22.0	11.1	31.0	14.2	54.5
South (HS)	33.8	15.4	16.9	8.5	28.5	16.2	20.8	9.2	50.8
West (HW)	36.1	15.4	12.3	7.3	23.8	10.1	27.8	13.7	54.6
Mid-Norway (HM)	31.7	11.2	13.7	6.8	25.5	18.0	29.2	7.5	56.5
North (HN)	45.6	14.4	8.9	1.1	24.4	11.1	21.1	8.9	64.4
Education level*,**									
Undergrad. and below	32.1	12.6	18.2	9.6	23.1	11.4	26.6	12.6	53.9
Graduate	36.8	14.4	9.6	4.1	26.4	14.4	27.2	10.5	56.6
Gross household income^{*,NS}									
Q1	31.5	13.6	17.5	8.6	19.8	9.7	31.2	14.5	53.5
Q2	35.1	13.2	10.7	5.9	30.7	19.0	23.4	10.2	51.7
Q3	35.2	12.2	12.2	5.9	26.6	12.5	26.0	9.9	59.5
Q4	38.9	15.9	10.6	4.4	25.7	13.7	24.8	9.7	56.2
Centrality class[b] ^{NS,NS}									
Class 1 (most urban)	36.9	13.7	13.3	7.3	24.9	12.9	24.9	10.7	55.4
Class 2	34.0	12.8	10.3	5.8	24.4	11.9	31.4	12.8	56.7
Class 3	34.9	13.4	15.4	5.5	26.4	13.7	23.3	11.6	55.8
Class 4	36.7	16.7	13.3	6.7	26.0	14.0	24.0	8.0	54.7
Class 5	23.4	10.4	20.8	10.4	20.8	15.6	35.1	13.0	50.6
Class 6 (most rural)	43.3	16.7	6.7	6.7	23.3	10.8	26.7	13.3	53.3

Notes:

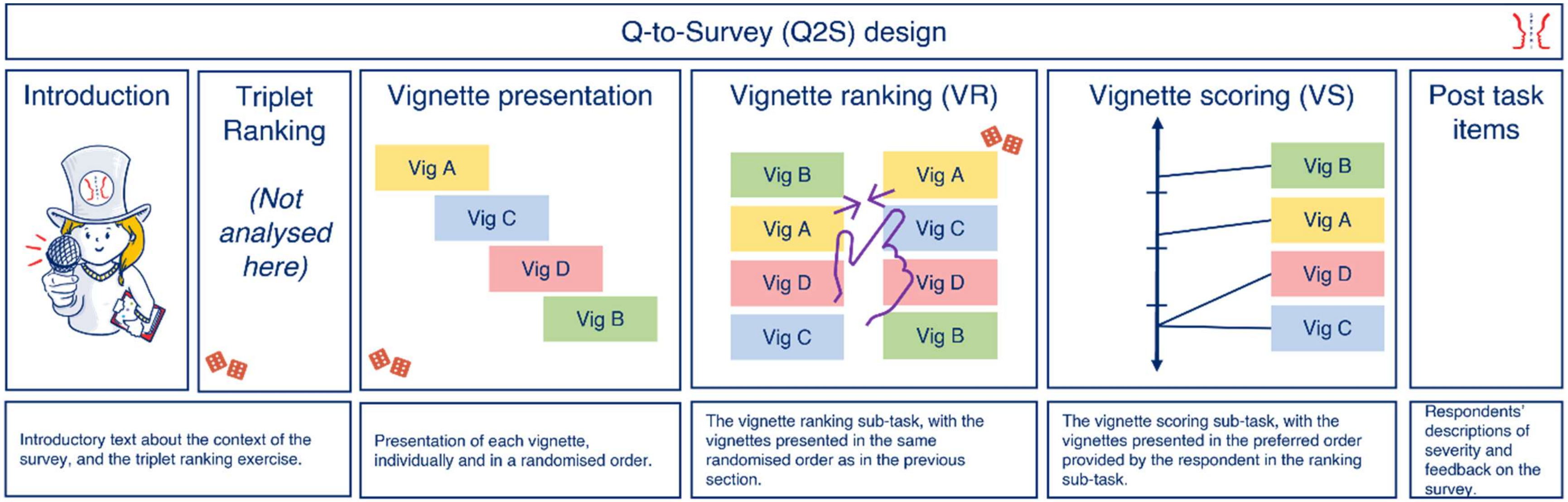
^{NS} $p \geq 0.05$, $*\,p < 0.05$, $**\ p < 0.01$, $***\ p < 0.001$ for χ^2 test for equality of distributions with respect to C3 and C4, respectively for each of the sociodemographic variables.

[a] *HM,* Health Mid; *HN*, Health North; *HS*, Health South; *HW*, Health West. Norway's health trusts are divided into four health regions: South-East, corresponding to the Norstat regions Oslo, East, and South, and the remaining three (Middle, West, and North) corresponding to the Norstat-defined regions.

[b] The target proportions for centrality classes were estimated from Table 5.2 in Statistics Norway (Høydahl, 2020).

Figures

Figure 1 Overview of survey design.



Introductory text about the context of the survey, and the triplet ranking exercise.

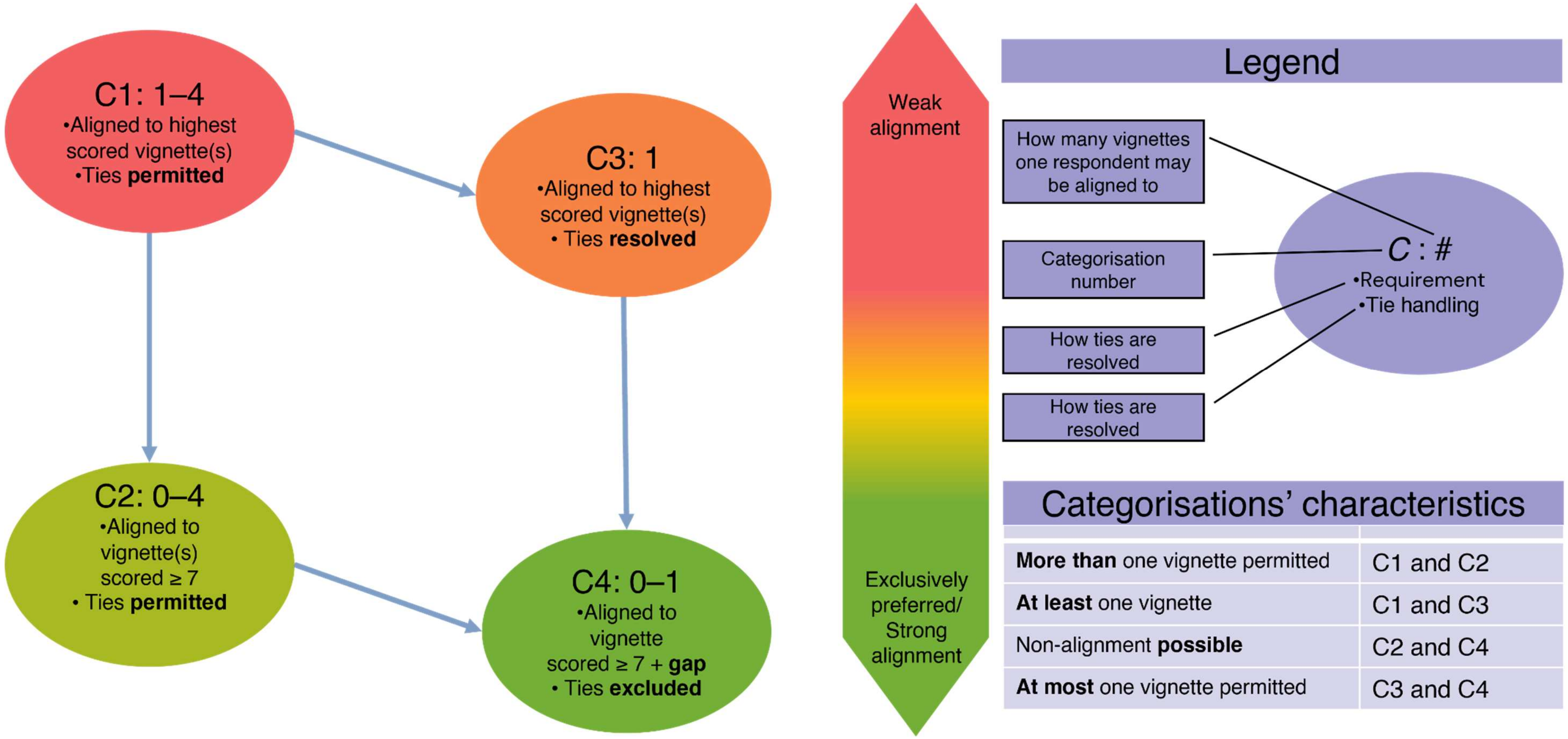
Presentation of each vignette, individually and in a randomised order.

The vignette ranking sub-task, with the vignettes presented in the same randomised order as in the previous section.

The vignette scoring sub-task, with the vignettes presented in the preferred order provided by the respondent in the ranking sub-task.

Respondents' descriptions of severity and feedback on the survey.

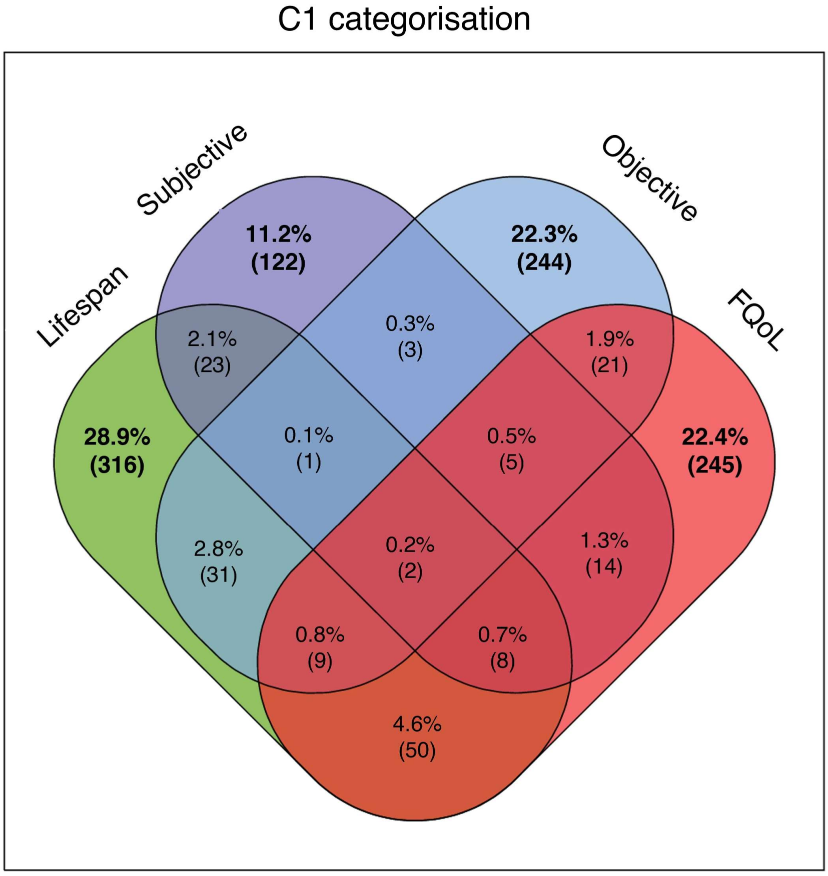
Figure 2 Interdependencies between the categorisations (C1 to C4).



Categorisations' characteristics

More than one vignette permitted	C1 and C2
At least one vignette	C1 and C3
Non-alignment **possible**	C2 and C4
At most one vignette permitted	C3 and C4

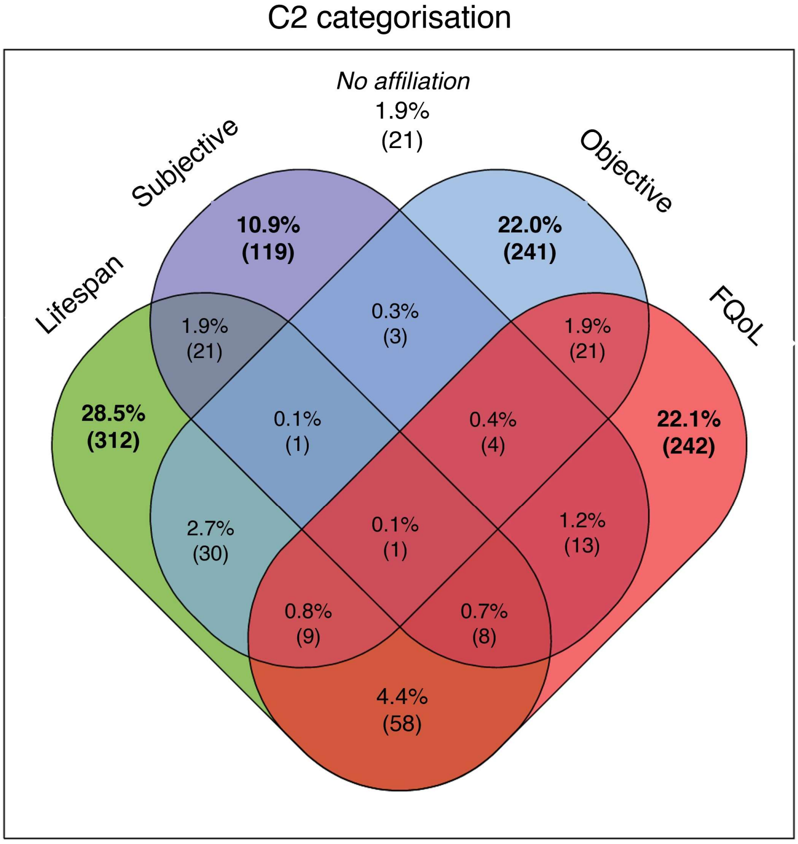
Figure 3 Venn diagram representation of the results of the C1 analysis. Numbers in brackets represent the number of respondents classified as aligned with the vignette(s) in the C1 categorisation[a].



FQoL, Functioning and Quality of Life.

[a] Bold text is used to illustrate respondents who by C1 criteria are placed only on one vignette.

Figure 4 Venn diagram representation of the results of the C2 analysis. Numbers in brackets represent the number of respondents classified as aligned with the vignette(s) in the C2 categorisation[a].



FQoL, Functioning and Quality of Life

[a] Bold text is used to illustrate respondents who by C2 criteria are placed only on one vignette.

References

Abelson, J., Blacksher, E. A., Li, K. K., Boesveld, S. E., & Goold, S. D. (2013). Public Deliberation in Health Policy and Bioethics: Mapping an emerging, interdisciplinary field. *Journal of Deliberative Democracy*, *9*(1), Article 1. https://doi.org/10.16997/jdd.157

Backor, K., Golde, S., & Nie, N. (2007). Estimating Survey Fatigue in Time Use Study. *The Stanford Institute for the Quantitative Study of Society (SIQSS)*. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=401f97f2d7c684b295486636d8a84c627eb33446

Baker, R., Mason, H., McHugh, N., & Donaldson, C. (2021). Public values and plurality in health priority setting: What to do when people disagree and why we should care about reasons as well as choices. *Social Science & Medicine*, *277*, 113892. https://doi.org/10.1016/j.socscimed.2021.113892

Baker, R., van Exel, Job, Mason, Helen, & Stricklin, Michael. (2010). Connecting Q & Surveys: Three Methods to Explore Factor Membership in Large Samples. *Operant Subjectivity: The International Journal of Q Methodology*, *34(1)*, 38–58.

Barra, M., Broqvist, M., Gustavsson, E., Henriksson, M., Juth, N., Sandman, L., & Solberg, C. T. (2019). Severity as a Priority Setting Criterion: Setting a Challenging Research Agenda. *Health Care Analysis*, 1–20. https://doi.org/10.1007/s10728-019-00371-z

Bonaccio, M., Di Castelnuovo, A., de Gaetano, G., & Iacoviello, L. (2020). Socioeconomic gradient in health: Mind the gap in 'invisible' disparities. *Annals of Translational Medicine*, *8*(18), 1200. https://doi.org/10.21037/atm.2020.04.46

Broqvist, M. (2018). *Asking the public: Citizens´ views on priority setting and resource allocation in democratically governed healthcare*. https://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-146217

Broqvist, M., Sandman, L., Garpenby, P., & Krevers, B. (2018). The meaning of severity – do citizenś views correspond to a severity framework based on ethical principles for priority setting? *Health Policy*, *122*(6), 630–637. https://doi.org/10.1016/j.healthpol.2018.04.005

Brown, S. R. (1993). A Primer on Q Methodology. *Operant Subjectivity*, *16*.

Brown, S. R. (2002). Q technique and questionnaires. *Operant Subjectivity*, *25*, 117–126.

Chalkidou, K. (2012). Evidence and values: Paying for end-of-life drugs in the British NHS. *Health Economics, Policy, and Law*, *7*(4), 393–409. https://doi.org/10.1017/S1744133112000205

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford University Press.

Flick, U. (2013). *The SAGE Handbook of Qualitative Data Analysis*. SAGE.

Gu, Y., Lancsar, E., Ghijben, P., Butler, J. R., & Donaldson, C. (2015). Attributes and weights in health care priority setting: A systematic review of what counts and to what extent. *Social Science & Medicine*, *146*, 41–52. https://doi.org/10.1016/j.socscimed.2015.10.005

Guest, G., M.MacQueen, K., & E.Namey, E. (2012). *Applied Thematic Analysis*. SAGE Publications, Inc. https://doi.org/10.4135/9781483384436

Hartmannsgruber, M. W. B., & Silverman, M. D. G. (2000). Applying Parametric Tests to Visual Analog Scores. *Anesthesia & Analgesia*, *91*(1), 248. https://doi.org/10.1213/00000539-200007000-00059

Hausman, D. (2019). The significance of 'severity'. *Journal of Medical Ethics*, *45*(8), 545–551. https://doi.org/10.1136/medethics-2018-105058

Helseforskningsloven. (2008). *Lov om medisinsk og helsefaglig forskning (helseforskningsloven)—Lovdata*. https://lovdata.no/dokument/NL/lov/2008-06-20-44

Høydahl, E. (2020). *Sentralitetsindeksen. Oppdatering med 2020-kommuner*. Statistics Norway. https://www.ssb.no/befolkning/artikler-og-publikasjoner/_attachment/413602?_ts=17085d29f50

Laguilles, J. S., Williams, E. A., & Saunders, D. B. (2011). Can Lottery Incentives Boost Web Survey Response Rates? Findings from Four Experiments. *Research in Higher Education*, *52*(5), 537–553. https://doi.org/10.1007/s11162-010-9203-2

Lønning, I., Album, D., Benkow, J., Brofoss, K., Brundtland, G., Buttedahl, A., Evensen, S., Nesheim, B.-G., Sandmo, A., Traaholt, I., Shetelig, A. W., Wislæff, N. F., Ødegård, L., Aase, K., Halvorsen, M., Husbyn, H., & Norheim, O. F. (1997). *Prioritering på ny—Gjennomgang av retningslinjer for piroriteringer innen norsk helsetjeneste* (NOU 1997:18; Norges offentlige utredninger, p. 225). Sosial- og helsedepartementet.

Lønning, I., Asbøll, E., Berge, T. E., Evensen, Å. R., Grund, J., Gunbjørud, A. B., Holm, N., Kristiansen, E., Normann, T., Rasmussen, K., Thu, K., Tranøy, K. E., Myhre, L., & Brofoss, K. (1987). *Retningslinjer for prioriteringer innen norsk helsetjeneste* (NOU 1987:23; Norges offentlige utredninger, p. 108). Sosialdepartementet.

Magnussen, J., Aaserud, M., Granaas, T., Magelssen, M., Syse, A., Celius, E. G., Klovning, A., & Syversen, I. D. (2015). *På ramme alvor—Alvorlighet og prioritering* (p. 71) [Rapport fra arbeidsgruppe nedsatt av HOD]. Department of Health.

Marmot, M. (2004). Status Syndrome. *Significance*, *1*(4), 150–154. https://doi.org/10.1111/j.1740-9713.2004.00058.x

Mason, H., Collins, M., McHugh, N., Godwin, J., Van Exel, J., Donaldson, C., & Baker, R. (2018). 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 Economics*, *27*(5), 819–831. https://doi.org/10.1002/hec.3640

Mason, H., van Exel, J., Baker, R., Brouwer, W., Donaldson, C., Donaldson, C., Baker, R., Mason, H., Pennington, M., Bell, S., Jones-Lee, M., Wildman, J., Lancsar, E., Robinson, A., Bacon, P., Olsen, J. A., Gyrd-Hansen, D., Kjaer, T., Beck, M., … Shackley, P. (2016). From representing views to representativeness of views: Illustrating a new (Q2S) approach in the context of health care priority setting in nine European countries. *Social Science & Medicine*, *166*, 205–213. https://doi.org/10.1016/j.socscimed.2016.08.036

Ministry of Health and Care Services. (2020). *Meld. St. 38 (2020-2021) Nytte, ressurs og alvorlighet—Prioritering i helse- og omsorgstjenesten*. Helse- og omsorgsdepartmentet. https://www.regjeringen.no/contentassets/85eb048de42342e4b789913cb8b1b218/no/pdfs/stm202020210038000dddpdfs.pdf

Myles, P. S., Troedel, S., Boquest, M., & Reeves, M. (1999). The pain visual analog scale: Is it linear or nonlinear? *Anesthesia and Analgesia*, *89*(6), 1517–1520. https://doi.org/10.1097/00000539-199912000-00038

National Institute for Health and Care Excellence. (2022). *NICE health technology evaluations: The manual*. 181.

Nord, E., & Johansen, R. (2014). 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*, *116*(2), 281–288. https://doi.org/10.1016/j.healthpol.2014.02.009

Norheim, O. F. (2010). Priority To The Young Or To Those With Least Lifetime Health? *The American Journal of Bioethics*, *10*(4), 60–61. https://doi.org/10.1080/15265161003697305

Norheim, O. F., Allgott, B., Aschim, B., Førde, R., Gjul, G. K., Gundersen, T., Kakad, M., Kjellevold, A., Kvinsland, S., Melberg, H. O., Moen, A., Mæland, Ø., Olsen, J. A., & Sjøli, S. (2014). *Åpent og rettferdig—Prioriteringer i helsetjenesten* (NOU 2014:12; p. 219). Helse- og

Omsorgsdepartementet. https://www.regjeringen.no/no/dokumenter/NOU-2014-12/id2076730/

Norstat. (2023). *Privacy Policy | Norstatpanel*. https://www.norstatpanel.com/en/privacy-policy

Olsen, J. A. (2013). Hva menes med «sykdommens alvorlighetsgrad»? *Tidsskrift for Den norske legeforening*. https://doi.org/10.4045/tidsskr.12.0659

Ottersen, T., Mæstad, O., & Norheim, O. (2014). Lifetime QALY prioritarianism in priority setting: Quantification of the inherent trade-off. *Cost Effectiveness and Resource Allocation*, *12*(1), 2. https://doi.org/10.1186/1478-7547-12-2

Reckers-Droog, V. T., van Exel, N. J. A., & Brouwer, W. B. F. (2018). Looking back and moving forward: On the application of proportional shortfall in healthcare priority setting in the Netherlands. *Health Policy*, *122*(6), 621–629. https://doi.org/10.1016/j.healthpol.2018.04.001

Reynolds, J. M. (2022). *The Life Worth Living*. University of Minnesota Press.

Riksdagsförvaltningen. (1997). *Prioriteringar inom hälso- och sjukvården Socialutskottets Betänkande 1996/97:SoU14—Riksdagen*. https://www.riksdagen.se/sv/dokument-lagar/arende/betankande/prioriteringar-inom-halso--och-sjukvarden_GK01SoU14

Riksdagsförvaltningen. (2018, November 21). *Prioriteringar inom hälso- och sjukvården Kommittédirektiv 1992:8—Riksdagen*. https://www.riksdagen.se/sv/dokument-lagar/dokument/kommittedirektiv/prioriteringar-inom-halso--och-sjukvarden_GGB18

RStudio Team. (2020). *Integrated Development for R* [RStudio]. http://rstudio.com/

Schurer, M., Matthijsse, S. M., Vossen, C. Y., van Keep, M., Horscroft, J., Chapman, A.-M., & Akehurst, R. L. (2022). Varying Willingness to Pay Based on Severity of Illness: Impact on Health Technology Assessment Outcomes of Inpatient and Outpatient Drug Therapies in The Netherlands. *Value in Health*, *25*(1), 91–103. https://doi.org/10.1016/j.jval.2021.08.003

Shah, K. K. (2009). Severity of illness and priority setting in healthcare: A review of the literature. *Health Policy*, *93*(2/3), 77–84.

Skedgel, C., Henderson, N., Towse, A., Mott, D., & Green, C. (2022). Considering Severity in Health Technology Assessment: Can We Do Better? *Value in Health*, *25*(8), 1399–1403. https://doi.org/10.1016/j.jval.2022.02.004

Stenmarck, M. S., Jølstad, B., Baker, R., Whitehurst, D. GT., & Barra, M. (2023). A severely fragmented concept: Uncovering citizens' subjective accounts of severity of illness. *Social Science & Medicine*, *330*, 116046. https://doi.org/10.1016/j.socscimed.2023.116046

Toombs, S. K. (1995). The lived experience of disability. *Human Studies*, *18*(1), 9–23. https://doi.org/10.1007/BF01322837

van Exel, N. J. A., Koolman, A. H. E., de Graaf, G., & Brouwer, W. B. F. (2011). *Overweight and obesity in Dutch adolescents: Associations with health lifestyle, personality, social context and future consequences*. Erasmus MC.

Watts, S., & Stenner, P. (2022, April 10). *Doing Q Methodological Research*. SAGE Publications Ltd. https://uk-sagepub-com.ezproxy.uio.no/en-gb/eur/book/doing-q-methodological-research

Williams, A. (1997). Intergenerational Equity: An Exploration of the 'Fair Innings' Argument. *Health Economics*, *6*(2), 117–132. https://doi.org/10.1002/(SICI)1099-1050(199703)6:2<117::AID-HEC256>3.0.CO;2-B

Supplementary Material A (SMA): Vignette development

A Q2S study is based on a previously conducted Q study, and different methods can be applied in the design of the survey. The study presented here is a vignette-based study, where the viewpoints identified in the previous Q study (Stenmarck et al., 2023) are presented as short descriptions, intended to represent the central aspects of each viewpoint. Two central aspects when summarising the viewpoints into vignettes are salience and distinction. Salient features are those that respondents within the different viewpoints in the Q study feel strongly about, whereas distinguishing features are those that set the viewpoints apart (Mason et al., 2016). These two aspects were the main considerations when we developed the vignettes. The vignettes were initially drafted by MSS, then reviewed and edited in collaboration with all authors.

The viewpoints in the original Q study were based on statements collected from group interviews that were conducted in Norwegian. The vignettes were first written in Norwegian, to represent as accurately as possible the original meaning. They were then translated into English by MSS and reviewed by RB and DGTW. Both the Norwegian and English versions are presented in SMA Table 1.

SMA Table 1 Summary descriptions of the four viewpoints.

Norwegian language version	English language version
Vignette I: Lifespan (93 words)	
Jeg synes alvorlighet handler om hvordan helseproblemer påvirker den naturlige gangen gjennom livets ulike faser – og særlig hvis det påvirker unge. Døden er ikke nødvendigvis alvorlig, spesielt når du er gammel, og det finnes ting som kan være mer alvorlig enn døden. Smerter trenger heller ikke være alvorlig. Men ting som psykisk sykdom og tap av verdighet, det er alvorlig. Alvorlighet er altså først og fremst når sykdom rammer den naturlige gangen gjennom livets ulike faser, og fratar noen muligheten til å oppleve det man burde få lov til å forvente av livet.	I think severity is about how health problems affect the natural course of life, and affect the natural development of life – and especially if it affects the young. Death isn't necessarily severe, especially when you're old, and there are things that are more severe than death. Pain isn't necessarily severe either. But things like mental illness and loss of dignity, that is severe. Severity is first and foremost when illness affects the natural course of life's different phases, and takes one's possibility to experience what you should be allowed to expect from life.
Vignette II: Subjective (112 words)	
Jeg synes alvorlighet handler nesten bare om hvordan man selv opplever helseproblemene, og alvorlighet kan ikke knyttes til bestemte diagnoser eller tilstander. Alvorlighet avhenger helt av den opplevde situasjonen, og hva man føler er alvorlig for seg og sitt liv. Man kan ikke definere alvorlighet objektivt. Det er ikke noen fasit på hva alvorlighet er for noe, men man kan vel si at når man rammes i ung alder, eller det er noe som bare blir verre og verre, eller som fører til at du mister verdigheten, da blir det mer alvorlig. Kanskje det er mer alvorlig hvis det påvirker en som har foreldreansvar. Men det er vanskelig å si noe helt definitivt om alvorlighet.	I think severity is almost entirely about how one experiences health problems, and severity can't be tied to any specific diagnoses or conditions. Severity depends completely on the experienced situation, and what one feels is severe for oneself and one's life. You can't define severity objectively. There is no right answer to what severity is, but when you are affected at a young age, or by something that just gets worse and worse, or leads to a loss of dignity, that makes it more severe. And it's maybe more severe if it affects a parent who is responsible for a child. But it's hard to say anything definitive about what severity is.

Vignette III: Objective (91 words)

Jeg synes alvorlighet må defineres ut ifra noen objektive kriterier, som alder, diagnose, prognose, eller hastegrad. Det kan ikke være opp til den enkelte å bestemme hva som er alvorlig. Vi trenger noen kriterier, og helsepersonell eller andre eksperter er nok de beste til å uttale seg om hva som er alvorlig. Ting som psykisk sykdom, smerter og død kan være alvorlig, eller det å ikke kunne jobbe. Så, alvorlighet handler ikke om hva man selv tror, føler, eller synes. Hvor alvorlig noe er burde bestemmes av objektive fakta om tilstander og diagnoser folk har.	I think severity has to be defined by some objective measures, like age, diagnosis, prognosis, and urgency. It can’t be up to each individual to decide what is severe and what isn’t. We need some criteria, and health personnel or other experts should be involved in making those criteria. Both mental illness and pain can be severe, or not being able to work, but severity is not about what someone believes, feels, or thinks. The degree of severity is decided by objective facts about the condition or diagnosis that one has.

Vignette IV: Functioning and Quality of Life (80 words)

Jeg synes alvorlighet avhenger av hvordan det påvirker hverdagslivet ditt: om det endrer funksjonsevnen din på jobb eller hjemme, eller evnen din til å delta i samfunnet, drive med hobbyene dine og sånne ting. Likevel kan ikke hva som helst være alvorlig. Det er helt klart en slags objektiv sammenheng mellom helseproblemer (psykiske eller fysiske) og hvordan livskvaliteten blir påvirket. Men alvorlighet bør måles gjennom hvordan helseproblemet påvirker livskvaliteten og funksjonsnivået ditt.	I think severity depends on how it affects your day-to-day life: if it alters your functioning at work, at home, your ability to participate in society, enjoy your hobbies and things like that. Still, not just anything can be severe. There are clearly some objective links between a health problem (mental or physical) and how your quality of life is affected. But severity has to be measured by how it affects your quality of life and your levels of functioning.

Supplementary Material B (SMB): Survey feedback

In the final section of the survey, respondents were asked to rate how difficult the ranking and scoring subtasks were to complete, on a scale from 1 (‘very difficult’) to 5 (‘very easy’). Only response options 1 and 5 were labelled; there was also a ‘don’t know’ option for each subtask. Using a similar format, participants were asked how understandable the instructions in the survey were (1, ‘very poor/difficult to understand’; 5, ‘very good/easy to understand’). Results for these three questions are presented in SMB Table 1. More than 95% of respondents gave a lower score (indicating more difficulty) for the questions about the difficulty of the tasks when compared with the question about the quality of the survey instructions.

SMB Table 1 Respondents’ feedback on the difficulty of the ranking and scoring tasks, and the quality of instructions. Values are numbers (percentages).

Rating	Ranking subtask	Scoring subtask	Survey instructions
1 (highest level of difficulty)	106 (9.7)	107 (9.8)	15 (1.4)
2	324 (29.6)	290 (26.5)	39 (3.6)
3	352 (32.2)	338 (30.9)	154 (14.1)
4	236 (21.6)	273 (25.0)	366 (33.5)
5 (lowest level of difficulty)	73 (6.7)	83 (7.6)	517 (47.3)
Don’t know[a]	3 (0.3)	3 (0.3)	3 (0.3)

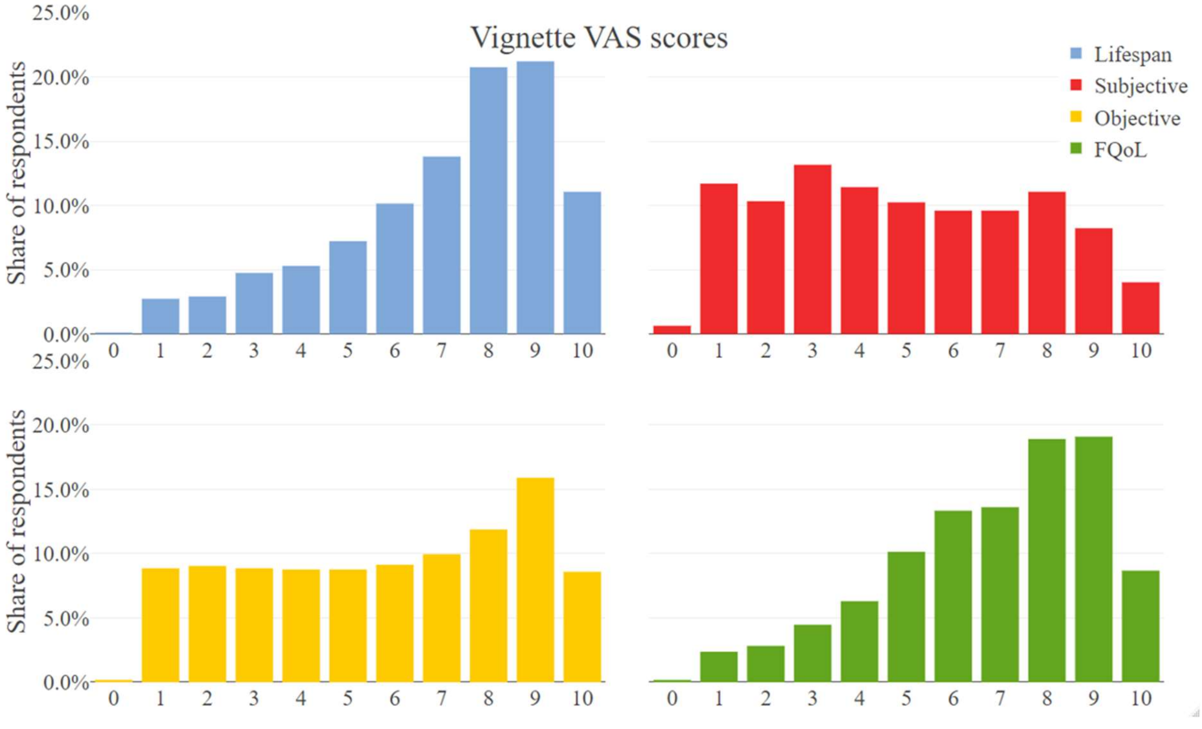
[a] The same three respondents reported ‘don’t know’ for the ranking and scoring subtasks. Three different respondents reported ‘don’t know’ regarding the quality of the survey instructions.

Supplementary Material C (SMC): A summary of (i) respondents' rearranging of randomly ordered vignettes, (ii) the use of the VAS scoring range, and (iii) descriptive statistics and correlation analysis for the VAS scores

We analysed how often the vignettes were rearranged from the randomly presented order in the ranking task. This was of interest because it is possible that the randomly presented order matched a participant's preferred ordering. There are 24 (4! [4x3x2x1]) possible ways to order the four vignettes. Assuming each respondent had a preferred ranking, one would expect 1 in 24 (4.2%) participants to not rearrange the randomly ordered vignettes presented to them. We found that 6.9% of participants did not rearrange the vignettes—a higher proportion than the expected 4.2%.

Respondents consistently scored the vignettes across the range of possible scores on the VAS. The average range (the difference between the highest score assigned to a vignette and the lowest score assigned to a vignette) was 6; 49.9% of respondents had a range between 5 and 7, with a further 25.1% having a range between 8 and 10. Respondents reported scores in the upper half of the 0-10 scale more often than the lower half, with 62.2% of scores equal to or greater than 6. The average highest score was 8.9 and the lowest maximal score was 3. The distributions of VAS scores for the four vignettes are shown in SMC Figure 1.

SMC Figure 1 The distributions of VAS scores (x axes) for the four vignettes.



Summary statistics and proportions across the trisected categories for the VAS scores are reported, by vignette, in SMC Table 1. Of the four vignettes, Lifespan had the highest mean and median VAS scores, as well the highest proportion of reported alignment (i.e., a VAS score equal to or greater than 7). The FQoL vignette had the second highest mean, median, and proportion of reported alignment, with the Objective vignette in third for all three statistics. Approximately 10% of respondents expressed no alignment (i.e., a VAS score of three or less) with the Lifespan and FQoL vignettes. The Subjective vignette had the lowest mean and median VAS scores and a relatively even distribution across the three alignment categories; the highest proportion of respondents indicating ‘no affiliation’ was for the Subjective vignette. Overall, 1.1% of VAS scores were 0, while 32.4% of VAS scores were 10.

SMC Table 1 Summary statistics and trisected categories for the VAS scores, by vignette.

Vignette	Median	Mean (SD)	Alignment (7-10)	Neutral (4-6)	No alignment (0-3)
Lifespan	8.0	7.1 (2.3)	66.8%	22.7%	10.5%
Subjective	5.0	5.0 (2.7)	32.9%	31.3%	35.8%
Objective	6.0	5.8 (2.9)	46.3%	26.7%	27.0%
FQoL	7.0	6.8 (2.3)	60.3%	29.8%	9.9%

FQoL, Functioning and Quality of Life; *SD*, standard deviation

Pairwise comparisons of vignette VAS scores were further explored using correlation analysis (see SMC Table 2). At a 1% level of significance, the only non-significant pairwise correlations were the ones including the FQoL vignette. A statistically significant ($p<0.001$) negative correlation was observed between VAS scores for the Subjective and Objective vignettes.

SMC Table 2 Pearson correlation coefficients (p-value) for the pairwise comparisons of vignette VAS scores.

	Lifespan	Subjective	Objective	FQoL
Lifespan	-	0.01 (0.771)	-0.08 (0.009)	0.08 (0.011)
Subjective	-	-	-0.36 (<0.001)	0.03 (0.359)
Objective	-	-	-	-0.07 (0.020)
FQoL	-	-	-	-

FQoL, Functioning and Quality of Life