weaving necessity

contextualisation practices

for achieving robust health care coverage decisions

Tineke Kleinhout - Vliek

Weaving necessity

Contextualisation practices for achieving robust health care coverage decisions

Tineke Kleinhout-Vliek

Colofon

Weaving Necessity. Contextualisation practices for achieving robust health care coverage decisions
Tineke Kleinhout-Vliek
ISBN: 978-94-6361-462-7

Cover design by Bas Kleinhout
Layout and print by Optima Grafische Communicatie, Rotterdam

Weaving Necessity
Contextualisation practices for achieving robust health care coverage decisions

Noodzakelijkheid Weven
Contextualiseringspraktijken voor het tot stand brengen
van robuuste vergoedingsbesluiten

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. R.C.M.E. Engels

and in accordance with the decision of the Doctorate Board.
The public defence shall be held on

Thursday 5 November 2020 at 13:30 hrs

by

Trijntje Hendrika Vliek
born in Wierden

Erasmus University Rotterdam

Doctoral committee

Promotors	Prof.dr. A. Boer
	Prof.dr. A.A. de Bont
Other members	Prof.dr. R.M.P.M. Baltussen
	Prof.dr. W.B.F. Brouwer
	Prof.dr. D.M.J. Delnoij

Table of Contents

THANKS

Still, what I want in my life
is to be willing
to be dazzled–
to cast aside the weight of facts

and maybe even
to float a little
above this difficult world.
I want to believe I am looking

into the white fire of a great mystery.
I want to believe that the imperfections are nothing–
that the light is everything–that it is more than the sum
of each flawed blossom rising and fading. And I do.

Mary Oliver, "The Ponds."

At the end of this stretch of road and the beginning of this dissertation, my heartfelt thanks go to…

Antoinette and **Bert**, for mountains of generous, constructive feedback, for your patience with my shortcomings, and for always being concerned with process just as much as content. I've never even been remotely tempted to use the (private) nicknames that presented themselves to me fairly early on. I will miss our conversations too! Antoinette, for letting me apprentice under you. You are fiercely intelligent, but I especially admire your courage. Bert, for your kindess, your unwavering commitment to keeping it clear and, more importantly, getting it *right*.

At het Zorginstituut: **Jacqueline**, for clarifying countless things for me, both big and small, always with honesty and humour, and for sharing part of your extensive library of ZIN documents. **Pé Mullenders**, **Reinoud Doeschot**, and, naturally, all my **anonymous interviewees**, for helping me understand, map, and navigate this foreign territory.

Meindert Boysen, **Matthias Perleth**, **Raf Mertens**, and **Romke van der Veen**, for being just as curious about and committed to these matters as Bert.

Inez de Beaufort, for six words that I won't repeat here but that significantly shaped this dissertation's theoretical angle and methodology.

Renske Taks, for retrieving and analysing piles of documents during my first maternity leave.

Wichor Bramer, for his valuable help with the data for Chapter 1.

Tiago Moreira, for generous comments on an early draft of Chapter 4.

Roland, 'his' Governance group and the others attending the Monday lunches over the years: **Adina**, **Amalia**, **Andreea**, **Annemiek**, **Anneloes**, **Antoinette**, **Bert B**, **Bert de G**, **Bethany**, **Chiara**, **Dara**, **David**, **Diana**, **Elizabeth**, **Erna**, **Gijs**, **Hanna**, **Hester**, **Hugo**, **Ian**, **Iris**, **Jacqueline**, **Jannine**, **Jan-Willem**, **Jeroen**, **Jitse**, **Johanna**, **Jolien**, **Josje**, **Julia**, **Kim**, **Kor**, **Koray**, **Lieke**, **Leonoor**, **Maarten J**, **Maarten K**, **Marcello**, **Marianne**, **Marica**, **Marjolijn**, **Marthe**, **Martijn**, **Mathilde**, **Nienke**, **Oemar**, **Paul**, **Pauline, Paulus, Regianne**, **Richard**, **Rik**, **Robert**, **Sabrina**, **Suzanne K-B**, **Suzanne R**, **Tessa**, **Tim**, **Violet**, and **Wilma** (with my sincere apologies to anyone I have forgotten, I owe you a drink); for providing such sociable yet fertile ground – and for the annual Christmas speech joke to that effect.

Martijn, **Josje**, and **Dara**, for the most ineffective writing weekend ever. We had great conversations, **David** had a baby, all was well. All four, for all the halved burdens and doubled victories, scientific and otherwise. Josje, for your generosity of heart (I think I owe you eight birthday cards) and for our shared commitment to mixing languages. Dara, for never a dull moment: it's been pure caffeinated joy. Martijn, our professor, for relativism – and for not thanking me for that.

Bert de G, **Marianne**, **Martijn**, and **Marthe**, for great teamwork. In the words of one of my favourite novelists: let's plan a heist next.

Robert, for a couple of well-timed and enjoyable chats about much more than STS.

Matthijs Versteegh and **Vivian**, for friendly introductions to burden/severity of illness/disease.

Joeke and **Sanne**, for the opportunity of diving into the substandard & falsified medicines world together and for being so easy (and so fun!) to supervise. **Adina**, **Amalia**, **Elizabeth**, **Koray**, and **Maarten K**, for the inspiration on this topic and for a memorable picnic.

All the students I had the privilege of teaching, for many good conversations and insights.

Laura, for a fun start and for continuing support.

Bas, mijn 'broertje', for introducing me to Pantone colours.

Toos en **Gertjan**, for checking the Dutch summary but especially for the countless ways in which you help our family life run smoothly.

Evan en **Marjolein**, **Sjoerd** en **Renske**, **Jochem** en **Sanne**, for not letting me forget the lingo of my first scientific tribe, and for shared metaphorical and actual boats. Here's to the next decades of sailing together.

Roos, for a lifelong pact and, more recently, for a shared positioning towards STS – I never saw that one coming. You know this is not the reason I'm booking my room in the elderly peoples' home next to yours though.

All at Redeemer International Church, but **Elsi** en **PJ**, the rest of the eldership team, and my Life Group in particular, for being home away from Home.

Denise, for your generous friendship. For our shared love of all the words but especially those of Mr. Ishiguro. This trip to Scotland *is* happening, you have my, erm, word.

Maria, for being my favourite sister and partner in crime, for showing me how it's done, and for all the silly and serious chats.

Mam en **pap**, for instilling in me an apparently unshakeable belief that I could actually do this. Also, for all the prayers.

Andrea, for making me laugh so hard so often and for being even more intent on getting this thing finished than me.

Beatrice, the best deadline there ever was, for your starry-eyed smile.

My **Tom**, for keeping me sane and for accompanying me on a road of infinitely greater importance than the career ladder. For all the food, all the encouragement, and all the hard-won joy.

I | Necessity and health care coverage decisions

17th-century English statesman Oliver Cromwell held that it knows no law. As far as Greek tragedy poets Euripides and Sophocles were concerned, nothing had more strength. Even Ares, the god of war, could not battle it: necessity. These days, dictionaries speak of something required, unavoidable, indispensable, enforcing even. The philosophical definition runs along the same lines: "the principle according to which something *must be so*, by virtue either of logic or of natural law". As more contemporary idiom would have it, however, necessity is the mother of invention – you can even make a virtue out of it.

On this scale from enforcing to inspiring, many things have been considered a necessity: education, faith, fiction, happiness, music, and even private jets. As in the case of music, the person who deemed it so (in this instance, soul music legend Ray Charles) may add the nuance that it concerns a personal, rather than a general, necessity: "this is necessary for me". In the realm of health care, this dynamic rears its head too. Is a certain form of health care a necessity for one person – or does this also hold for more than one? Historically, medical doctors were responsible for making the decision for the individual patient, but in many countries, this has since been supplemented by general decisions made for more than one patient. These decisions on what is, and what is not, necessary health care taken on a more collective, often (sub) national level deal with the necessity of the health care technology generally. Examples of widely acknowledged *unnecessary* health care may seem relatively easy to come by. For forms of health care such as cosmetic surgery (Russell et al., 2014) and Viagra (Stolk et al., 2002), many people intuit that provision for all prospective recipients may not be a necessity. Then again, both erectile dysfunction and port-wine stains, the removal of which is often considered cosmetic and thus unnecessary, may cause impeded social functioning and severe psychosocial problems, which does indicate it may be a necessity for some patients. So what, if we think about it, should be the arguments pertaining to necessity to back up a collective health care coverage decision?

One of the few settings where this question is and has been tackled head-on is the Dutch National Health Care Institute (in Dutch: *Zorginstituut Nederland*, ZIN, in this text: the Institute). The Institute is responsible for delivering advice to the Minister of Health concerning which health care technologies should, and should not, be part of the Dutch basic benefits basket. This benefits basket specifies what Dutch citizens are entitled to through their collective health insurance, which is mandatory for all citizens. After receiving advice on the in- or exclusion of the health care technology from the benefits basket, the Minister of Health takes the final decision. The process for arriving at the advised decision at the Institute has on average, over the years, employed four criteria (effectiveness, cost-effectiveness, feasibility, and necessity), with these criteria representing questions that are to be answered in the final advice. For effectiveness, the question is: is this health care technology effective (enough)? Similarly, for cost-effectiveness: is this cost-effective enough? The necessity criterion likewise asks: is it necessary? This last question has also been phrased as: is a claim on the collective solidarity justified?

THIS DISSERTATION

This dissertation describes the outcomes of a study commissioned to operationalise the necessity criterion. The provenance of this project lies with the Institute. The project was part of a larger move towards more research substantiation and support for the Institute's work. This coincided with the set-up of the research network Health Technology Assessment Netherlands (in Dutch: *Academische Werkplaats Verzekerde Zorg*) to act as a "bridge between research and policy". This centre represents a collaboration between Erasmus University Rotterdam, Utrecht University, and the Institute and aims to support research designed together with policy advisors (Zorginstituut Nederland, 2020).

In 2014, a working group of Institute employees considered the necessity criterion likely to benefit from further operationalisation. This group had concerns about the extent to which the then-current version of the formalised necessity criterion reflected political and societal values pertinent to the case at hand. They noted that necessity was not a clear-cut criterion for making coverage decisions, despite multiple operationalisation and re-operationalisation steps that had been taken over the years. At the turn of the century, just over ten years after the term was first coined by the Dunning committee (Commissie-Dunning, 1991), scholars had already considered necessity a "difficult to measure, non-uniform unit" (Poley, 2002, p. 2313). Since then, multiple policy reports had dealt with the necessity question. Despite these efforts, and even though the formulation as a criterion did ensure that most, if not all, coverage advice documents contain a section 'necessity', further operationalisation was still deemed beneficial. This dissertation reports on the study that was done to fulfil this need.

A BRIEF HISTORY OF OPERATIONALISING NECESSITY

This section will provide a brief history of how necessity as a criterion has been operationalised thus far. This overview ranges from 1991 to 2013, which saw the publication of the last relevant policy report before the start of this project in March 2015. I distinguish two general strands of thought in these documents: first is the substance of the criterion, as several steps have been taken to 'operationalise' the necessity criterion, that is, to specify the exact question(s) this criterion should answer. Second, there is a recurring acknowledgement that necessity should be established in deliberation, concurrently with the other criteria and, possibly, other argumentations in the appraisal phase of coverage decisions. In health care coverage policy vernacular, the appraisal is treated as distinct from the establishment of the knowledge base through assessment of (a subset of) the knowledge types available, referred to as 'the assessment'. Such knowledge bases may range from randomised controlled trials to patient-reported outcome measures, which may include data harvested from social media as well as 'real-world data', but also live patient input (Kalf et al., 2018; Makady et al., 2017; Moes et al., 2016; Wiering et al., 2017). The appraisal provides a valuation of these knowledge bases, an exploration of additional pertinent factors, and a formulation of the

advice, and may or may not be set up as a distinct moment in time and space (Jansen et al., 2017; Oliver et al., 2004; Patera & Wild, 2014; Walley, 2007). With their focus on the appraisal as where necessity and its arguments should be brought together, these policy documents also highlight the importance of procedures to embed the necessity criterion in, in addition to the substance of the criterion.

In 1991, the term 'necessity criterion' was coined in the report '*Kiezen en delen; rapport van de commissie Keuzen in de zorg*', authored by the 'Dunning committee' (Commissie-Dunning, 1991). This report was written at the time of a reorganisation of health care more generally in the Netherlands (Helderman et al., 2014). The Dunning committee was the first to introduce a necessity criterion for making choices in health care. The report advocated the community approach (as opposed to the individual or medical-professional approach) to establishing necessity. This meant that necessary care was defined as all care that enabled, sustained, and where possible improved opportunities for individuals to share existence with other members of society. The core question of necessity was defined as what care would be considered necessary from a communal point of view, resulting in a potential 'ranking' of care according to level of necessity.

The Dunning committee specifically advocated the use of what they termed a funnel, which was to go down in history as 'Dunning's funnel' (in Dutch: *de trechter van Dunning*). This funnel contained four sieves (necessity, effectiveness, efficiency, and 'for own account' – the last of which will also become significant in due course). The idea was that forms of health care would either pass through all four sieves and end up in the basic benefits basket that was positioned underneath the funnel or get 'caught' in one of the sieves and therefore not be provided. Notably, this contrasts with the actual use of the four criteria in several examples that are given in the report. Helpful here is the case of in vitro fertilisation (IVF), where the Dunning committee was divided on the necessity, struggling, for example, with the question whether societal functioning was hampered by being childless. There were questions on financing as well in terms of own responsibility: IVF is expensive, but as expensive as adoption, which was (and is) not part of the benefits basket. All in all, the committee concluded, IVF is not highly necessary and should be low on the ranking. This showcases how criteria already in 1991 served less as a sieve and more like arguments to be weighed concertedly. In fact, as the committee poses later in the report, the goal of the whole exercise is to provide "arguments to base choices in health care on" (Commissie-Dunning, 1991, p. 109).

After its inception in the Dunning committee's report, the life of necessity as a formalised criterion was to last for nearly three decades. In this time, it went through several changes. Several reports by other government bodies followed Dunning's efforts, such as the *Contours of the Basic Health Benefit Package* report by the Health Council of the Netherlands (in Dutch: *Gezondheidsraad)*, which stated that package management should be based on both scientific and societal grounds and underlined the significance of following good procedures (Gezondheidsraad, 2003). This was resonated by the first of a duo of reports by the Council for Public Health and Care (in Dutch: *Raad voor de Volksgezondheid en Zorg*, RVZ, now *Raad voor de Volksgezondheid en Samenleving*, RVS). The first was the *Sensible and Durable Care* report (Raad voor de Volksgezondheid en Zorg, 2006),

published June 2006. This report was primarily concerned with the criteria for coverage decisions (whereas the next report, *Just and Durable Care* (Raad voor de Volksgezondheid en Zorg, 2007), dealt with the formalised procedures). The *Sensible and Durable Care* report noted the difficulties with operationalising necessity "not only in the Netherlands but elsewhere" (Raad voor de Volksgezondheid en Zorg, 2006, p. 6). It also names 'necessity/need for care' as a criterion, specifying that: "the higher this is, the sooner care would qualify to be paid for from collective means" (Raad voor de Volksgezondheid en Zorg, 2006, p. 6). This is operationalised as individual severity of illness only, specifying this as "the severity of illness for the 'average' individual as underwritten by society" (Raad voor de Volksgezondheid en Zorg, 2006, p. 15). The report also mentions the criterion 'justice/solidarity' – which is considered even more difficult to operationalise. Four potential concepts are given (egalitarianism, utilitarianism, rule of rescue, and the Maximin principle) but these are all "theoretical, without practical consequences." Usually, the report continues, the consequences come down to "justice based on need for care or based on equal access" (Raad voor de Volksgezondheid en Zorg, 2006, p. 20). This is considered to hang together with the 'solidarity' criterion, which is in turn related to the Dunning committee's 'for own account'. Here, it is important to note that collective funding is not necessary for those technologies patients can pay for themselves but that "necessary care must be given independent of financial carrying capacity" (Raad voor de Volksgezondheid en Zorg, 2006, p. 20). All in all, these considerations are related but it does not become clear how this should be worked out in practice. As to the working of criteria in general, the report notes that "filters or sieves do not work, because the world is not black-and-white" (Raad voor de Volksgezondheid en Zorg, 2006, p. 22). Accordingly, the report argues, these criteria must be weighed concomitantly. Notably, it names the appraisal phase as the place where non-quantifiable criteria feature, where a "*societal correction* on the technically achieved decision becomes possible" (Raad voor de Volksgezondheid en Zorg, 2006, p. 6, emphasis added). Moreover, the report states, if the two differ, the final choice needs to be justified explicitly.

Later that year, in December 2006, The Institute's predecessor, the *College of Health Care Insurances* (in Dutch: *College voor Zorgverzekeringen*, CVZ), was the first to mention package principles in their first *Package management in Practice* report (in Dutch: *Pakketbeheer in de Praktijk*, abbreviated PiP). Necessity was one such principle, which was defined as whether "the disease or required health care justify a claim on solidarity given the cultural context" (College voor Zorgverzekeringen, 2006). Notably, CVZ was also clear that these package principles and the specific criteria that underlay them should be weighed at the same time and without hierarchy – again in the strictest sense deviating from Dunning's sequential funnel (though not from how it was likely used in practice).

Necessity was operationalised in PiP1 as a combination of first, severity of illness and second, 'need for care', which were then to be combined with third, 'for own account': the costs of the intervention on the individual level (College voor Zorgverzekeringen, 2006, p. 36), previously visible in the Dunning report. These three are worked out further in the rest of the text (and the appendix also adds "public health argumentations, such as dangers to third parties" (College voor Zorgverzekeringen, 2006, p. appendix I)). The first element, severity of illness, was defined prefer-

ably quantitatively, in comparable units. The authors admit that this may not always be practical, or the data may not always be available, and this may therefore also be done qualitatively. The second element, need for care, was to "get a picture" of the "appeal to care" that this group of patients may do (College voor Zorgverzekeringen, 2006, p. 36). The third element, for own account, was part of the necessity criterion but to be used to value and nuance the necessity 'data' from the assessment phase during the appraisal phase.

> There is a severity of illness or need for care, and there is an adequate intervention, and despite this positive score, the judgement may still be that the costs do not justify a call on the social health care insurance. Because necessity will rarely be expressed as a resolute 'yes' or 'no', we weigh the question whether something can be for [the patient's] own account in the appraisal phase. (College voor Zorgverzekeringen, 2006, p. 42)

Regarding health care aids, the for own account question had been worked out further, to contain 'common use' and 'financial accessibility'.

> There are questions of 'common use aid', or 'substitution for a commonly used aid'. If it concerns a one-off purchase, with a long use, low costs, that does also not vary much from the provisions a regular citizen would have in their house, then the conclusion may be that the aid can be for own account [the patient can pay for it his or herself]. (College voor Zorgverzekeringen, 2006, p. 42)

In addition, 'financial accessibility', containing both contributions and/or accumulation of costs, were to be considered as part of necessity of insurance (see appendix 1 of the report). Overall, many elements of necessity were again specified, but the appraisal phase was given as the primary place where such elements would be brought together and weighed.

At the end of 2006, then, the necessity criterion was operationalised differently in different places. Most notable were the introduction of individual severity of illness, preferably quantified, and the specification of the Dunning committee's for own account as individual cost and/or common use. In addition, the reports suggest a host of different argumentations pertaining to solidarity, the cultural context, and public health. However, the difficulty of operationalising these is also commented on in several places. Moreover, we see an acknowledgement that criteria acting as filters or sieves does not work: criteria should be weighed concertedly and without hierarchy. In particular, the necessity question (the answer to which was considered to be rarely a terse yes or no) should be dealt with in an appraisal.

The following report of note was the second RVZ report, *Just and Durable Care* (Raad voor de Volksgezondheid en Zorg, 2007). It was, as far as I am aware, the first to propose the installation of a specific committee at CVZ: an appraisal committee. This report was followed by *Package management in Practice 2* (College voor Zorgverzekeringen, 2009). Necessity was defined the same as in

the first PiP report, but the criterion here comprised first, severity of illness, and second, necessity of insurance. The first element remained largely the same, but this second element deserves a closer look. It was defined as "whether it is necessary or due from a societal perspective to insure a health care intervention." The report continues,

> Experience teaches that it is difficult to judge during the assessment whether a health care intervention is necessary to insure. These considerations fit, after all, in the societal debate (appraisal). (College voor Zorgverzekeringen, 2009, p. 18)

Further on, the report states that the effectiveness and cost-effectiveness calculations will be weighed against the outcome on the rating on feasibility, the individual severity of illness, expressed as a percentage of individual loss of health for a patient due to this disease, and 'necessity of insurance'.

> This rating yields arguments in favour of or against incorporation into the [basic benefits] package. Arguments against incorporation will lower the chance of a positive advice; arguments in favour of incorporation into the package will heighten the chance of a positive advice. Examples of arguments in favour include: rarity of the disease (orphan indications) combined with a lack of alternative treatment options, informal care (a high level of informal care is given to the patient, which puts a high pressure on the environment of the patient), and risks for public health outside of the patient. Examples of arguments against incorporation into the package are: little overlap with the domain of health care, a high total budget impact, unsuitability of insurance due to high prevalence, and unsuitability of insurance due to high autonomy of the patient. The entirety of these argumentations determines the outcome of the advice. (College voor Zorgverzekeringen, 2009, p. 22)

In this report, the role of the appraisal committee (in Dutch: *Adviescommissie Pakket*, ACP) is also worked out for the first time by CVZ. This committee is specifically tasked with weighing the comments on the concept advice, making an inventory of the considerations, and determining the direction of the advice and priorities. Moreover, the report states, there should not be a sharp line between the assessment and appraisal phases.

By 2009, then, two significant steps had been taken. First, on the operationalisation of the necessity criterion side, we see the specification of Dunning's for own account in terms of insurance logic, where a combination of high individual cost and low risk indicates that *insuring* a health care technology makes sense. Second, on the procedural side, RVZ's suggestion to install an appraisal committee had been followed by CVZ.

Package management in Practice 3 was published another four years later (Zorginstituut Nederland, 2013). The funnel of Dunning is now named explicitly, with 'own risk and responsibility' as part of the necessity of insurance element of the necessity criterion (Zorginstituut Nederland, 2013,

p. 32). For the severity of illness element, the proportional shortfall method[1] was introduced, and the reasons why this was chosen: because the societal *opinions* on the distribution of health were now also considered of importance. Examples of such opinions may be that there may be no age discrimination, that people with an immediate need should be helped first, or that those with the worst health condition should receive most care (Zorginstituut Nederland, 2013, p. 33).

The Institute for Medical Technology & Assessment (iMTA) and the Institute for Health Policy & Management (in Dutch: *instituut Beleid en Management van de Gezondheidszorg*, iBMG) worked together with CVZ to develop the 'necessity of insurance' element, as in PiP2 it was noted to be in need of further operationalisation. Niëns and colleagues developed a '2x4 checklist', based on the Institute's assessment framework for medical care (in Dutch: *beoordelingskader hulpmiddelenzorg*):

Health insurance as an instrument

1. Is the intervention customary care?
2. Is the intervention foreseeable?
3. Might there be under-usage of an intervention if it is not insured?
4. Might there be over-usage of an intervention if it is insured (moral hazard)?

Financial accessibility

5. Does the intervention substitute for something that the majority of the population also uses?
6. Can the (additional) treatment costs be borne by the individual patient?
7. Can the patient expect relevant savings (offsetting the costs) due to the intervention?
8. Are treatment costs incurred only once or are they structural in character? (Niëns, 2014)

These questions, according to PiP3, were to be asked from an "insurance perspective" (Zorginstituut Nederland, 2013, p. 37). The first set of questions concerns the chance that something happens and the risk of moral hazard (over-usage), the second set considers the financial impact this may have for the individual. Questions 1 and 6 were expanded on in the report. For question 1, the report states: "This question is meant to delineate the insured care from the usual course of events in society." It specifies that if it is first, a generally customary provision, it will not be insured (such as a braille watch). Second, when it is customary care, it concerns the usual care that partners, inhabiting parents, or other house mates usually give one another. For question 6, the report notes

1 Proportional shortfall is a method currently in use in Dutch coverage decision-making practice to quantify the 'necessity of care' for a certain health care technology. Its objective is to create more equity in terms of severity of illness and (prospective or past) health than counting all Quality-adjusted Life Years (QALYs) as equal, as the latter does not reflect wider (societal) notions in terms of treating those in greatest need first. It seeks to do so through quantifying the proportion of QALYs patients lost due to the disease without the technology compared to the remaining QALYs these patients would have had without this disease (Reckers-Droog et al., 2018; Stolk et al., 2004). It is currently used in practice in setting reference values for cost-effectiveness thresholds (Reckers-Droog et al., 2019; Zorginstituut Nederland, 2018).

that it has been difficult in practice to state a maximum amount, and that this is perhaps even undesirable. It also holds that "whether certain package proposals exceed the capacity of citizens is, ultimately, a political consideration. From insurance theory, it is a legitimate question whether the costs of provision outweigh the costs of insurance." The report argues that for low costs, insurance is not indicated, because if it is not provided through the benefits package "the market will do its job, in the good sense of the word" (Zorginstituut Nederland, 2013, p. 38). This, the report continues, should be regarded in cohesion with other considerations.

PiP3 thus further operationalised the necessity criterion's first element, severity of illness, as calculated by means of the proportional shortfall method, and the second element, necessity of insurance, as a 2x4 checklist. Notably, these highly specified considerations should still be considered in cohesion with other argumentations but not much attention is given in PiP3 as to how this should take place.

Summarising, the two strands of thought that have dominated these policy reports have both evolved over the years. The substantive operationalisation of necessity has moved from relatively broadly defined to a high degree of specification in the shape of a complex calculation and a checklist. Over the years, the specified criterion has contained several elements, the most transient of which seems to have been solidarity and public health argumentations, though an acknowledgement of the context, be it cultural or otherwise, appears regularly. Second, there has been an almost continual appreciation of deliberative settings for the establishment of necessity so as to be able to weigh it concurrently with other criteria (and other argumentations), with necessity relatively quickly losing its primacy as the first 'sieve' in Dunning's funnel. Establishment of necessity or answering the necessity question, in particular, has over the years been specified as best taking place in the appraisal phase of decisions by a separate appraisal committee.

DUTCH COVERAGE DECISION-MAKING PRACTICE

Current Dutch coverage decision practice, follows the generic assessment-appraisal pattern outlined above, with a few additional steps. After a form of health care has been set on the agenda as a topic requiring consideration from the Institute, a scoping session is initiated, in which stakeholders are invited to contribute relevant considerations. Next, a wide variety of knowledge types is gathered by Institute employees, which are subsequently assessed by the assessment committee (in Dutch: *Wetenschappelijke Adviesraad*, WAR). This is written up into a 'discussion document' which features headings per criterion and other relevant considerations to benefit the appraisal committee (in Dutch: *Adviescommissie Pakket*, ACP). The appraisal aims for an explicit "societal weighing" of the knowledge established in the assessment (art. 14, Zorginstituut Nederland, 2016, 2017), resulting in an advised decision. This advised decision is discussed and approved by the Institute's board of directors and sent to the Minister.

To arrive at this advice, a combination of both Health Technology Assessment (HTA) and Accountability for Reasonableness (A4R) is reportedly used (Zorginstituut Nederland, 2017). HTA and A4R are both institutionalised frameworks aiming to benefit coverage decision-making practice and both have been extensively refined and their workings in practice studied by scientists. They are relevant as the two strands of thought I distinguished in the policy documents above have each been resonated by, and probably influenced by, these two scientific fields. On the one hand, the idea of providing well-established, consistently-applied substantive criteria for making decisions has been advocated by scholars in the field of Health Technology Assessment. Scholars working on Accountability for Reasonableness, on the other hand, would consider good procedures the primary guarantee for good decisions. Below, I give a brief introduction to both[2].

HTA aims to provide a "systematic evaluation of the properties and effects of a health technology" (INAHTA, 2020). It does so assuming that the total budget a country has available to spend on health care technologies is limited and it seeks to identify which technologies provide value for money as a consequence (Lehoux, 2014). HTA was at its inception defined as the evaluation of both the technical side of a health care technology as well as the societal impact, that is, both assessment and appraisal. In practice, HTA has largely been narrowed down to effectiveness and cost-effectiveness calculations to benefit the assessment phase (Giacomini, 1999; Lehoux & Blume, 2000), despite historic and recent calls to (re)integrate 'ethical issues' into HTA (Daniels et al., 2016; Jansen et al., 2017). HTA remains a hugely impactful enterprise and these technical assessments are now more and more achieved internationally (Guegan et al., 2014; Stolk et al., 2009). A large part of the work of the Institute currently entails achieving an HTA per health care technology, which feeds into decision making through the assessment phase in particular and serves to 'ground' the advice scientifically (Niezen, 2012). This grounding is directly based on notions of evidence-based medicine (Abrishami, 2017; Lehoux & Blume, 2000).

On the procedural side, Daniels and Sabin developed the Accountability for Reasonableness (A4R) framework in the 1990s based on Rawls' theory of justice as fairness. A4R responded to questions around the legitimacy and fairness of limit-setting decisions made by USA-based Managed Care Organizations and other insurers at the time. A4R sets procedural boundaries for coverage decisions: relevance, publicity, appeals, and enforcement (Daniels, 2000; Daniels & Sabin, 1997, 1998, 2008). These respectively aim to ensure that 1) the rationale for the decision is supported by reasons judged reasonable by "fair-minded people"; 2) this rationale is provided

2 Recent efforts have attempted to bring the two together into a comprehensive framework for 'evidence-informed deliberative processes' (EDPs) (Bærøe & Baltussen, 2014; Baltussen et al., 2016; Baltussen et al., 2017; Oortwijn et al., 2020). Such EDPs should, these authors argue, "learn about the relevant social values" through early stakeholder involvement and evaluate these values in a manner that is informed by evidence. They denote five important implications for the ideal organisation of the processes at HTA agencies. First, these agencies should organise stakeholder involvement well. Second, they should integrate the assessment and appraisal phases. Third, it is suggested that the criteria for making decisions are subjected to public scrutiny. Fourth, HTA agencies are advised to formulate and use a checklist of criteria that are considered potentially relevant and, in the justification or rationale for decisions, outline how each criterion affected this decision. Fifth, the authors suggest making these justifications or rationales public and enabling appeals.

publicly; 3) this decision can be appealed; and 4) the first three criteria are enforced in some way. A4R thus prescribes criteria for coverage decision processes and holds that through satisfaction of these criteria, the process can be considered legitimate. A4R has been considered both acceptable and applicable in a wide variety of decision settings (Daniels & Sabin, 2008; Kapiriri et al., 2009; Kapiriri & Razavi, 2017; Martin et al., 2002). In essence, A4R holds that a decision outcome should be accepted as fair when it has been reached by a fair procedure (Daniels et al., 2016). As opposed to HTA, A4R thus sets procedures over content in terms of achieving health care coverage decisions. A recent policy report specifies how A4R's procedural boundaries are adhered to in practice at the Institute (Zorginstituut Nederland, 2017).

Of the whole coverage decision-making process, this dissertation will focus on the appraisal phase as this is where necessity is located according to the policy documents studied. Specifically, this dissertation will build on descriptive, inductive understandings of what appraisal entails. This will mean not situating it in either the HTA or A4R scientific tradition, as these are in essence both more prescriptive than descriptive, deriving as they do from larger principles for good decision making, namely justice as consistently-applied criteria and justice as fairness respectively. This means that these schools of thought are not primarily concerned with describing what happens in practice but with to what extent practice may adhere to the principles set out. Instead, I will take my cues from inductive studies which understand appraisal to be a deliberative process to come to a decision through interpretation of the knowledge input. These studies describe appraisals as featuring a plethora of different argumentations (Dakin et al., 2015; Franken et al., 2015; Gordon, 2006; Guindo et al., 2012; Jansen et al., 2017; Morrell et al., 2017; Shah, 2009), which may contradict (Martin et al., 2001; Singer et al., 2000), and the applicability of which may be under discussion (Kapiriri et al., 2009; Vuorenkoski et al., 2008). Two schools of thought have considered inductively how these types of decisions are arrived at: Health Services Research and Science and Technology Studies. I will discuss both in turn.

HEALTH SERVICES RESEARCH: ELEGANCE, RATIONALITY, EXPERTISE

This overview on the inductive work on coverage decision making in the Health Services Research (HSR) field starts in the middle of the 1990s. 'Muddling through', a key phrase in this field at this time, denoted decisions made 'not according to general rules', that is, decision criteria. Instead, decisions relied on the discretion of the decision makers (Hunter, 1995; Mechanic, 1997). The term muddling through is borrowed from the political scientist Lindblom (1959) who observed in his landmark study that even when you agree on both the (knowledge) input and the intended goals of a policy, you may still disagree on which policy is most appropriate. Preferences of decision makers vary, and they, therefore, value certain policies or policy instruments differently. This, however, Lindblom considered not so much a problem as a fact of policy life. Since its inception, muddling through has been criticised for resulting in inconsistency between decisions and even in

arbitrariness. Scholars in the coverage decision-making field, however, preferred it over the rigid application of universal decision criteria, which they considered to lead to too little attention for the case at hand (Entwistle et al., 1996; Ham, 1999). They described 'implicit' or individual-level decisions that relied on "discretion, flexibility, and ability to take account of emotions, aspirations, and preferences" (Mechanic, 1997, p. 90). Their answer to the arbitrariness charge was to muddle through *elegantly* with greater transparency concerning the grounds for a decision (Hunter, 1995; Mechanic, 1997). Much of the inductive work that followed this positioning described decision makers' subjectivities, that is, personal preferences and ways of deciding, showcasing "the practical circumstances of real-world decision making" (Hughes & Light, 2002, p. 1).

Since then, the field has moved towards describing the rationality of decisions, specifically favouring a both-and conceptualisation of rational decisions (Calnan et al., 2017; Gkeredakis et al., 2011; Hughes & Doheny, 2011). For rational decisions, decision makers not only take into account "contexts and occasions" but are accountable to criteria, such as those provided by evidence-based medicine, at the same time (Jenkings & Barber, 2004, p. 1765). The application of criteria and the incorporation of scientific knowledge has thus become part and parcel of making rational decisions, rather than being opposed to it. This has required decision makers to negotiate between and ultimately incorporate both those more formal rationalities and more 'local' ones in the decision-making process (Hughes & Doheny, 2011). This has resulted in an overall focus on the work that goes into combining; some scholars describe decision makers as being both 'rational' and 'human' (Russell & Greenhalgh, 2014); others how they display a "combining of strategies" (Calnan et al., 2017; Moreira, 2005). Such combining work has specifically been characterised as being *pragmatic* (Calnan et al., 2017; Hunter, 1995; Russell & Greenhalgh, 2014); resulting in decisions with a 'pragmatic rationality' (Russell, 2017, following a.o. Aristotle). Such rationality is juxtaposed with both instrumental rationality, where a complete set of decision rules guarantees the quality of the decision (embodied by HTA efforts), and institutional rationality, with transparency and good procedure acting as guarantors (as visible in the A4R framework) (Gkeredakis et al., 2011; Ham & Glenn, 2003; Russell, 2017). Pragmatic rationality carries a distinctly positive valuation as "a characteristic of expert judgement" (Russell, 2017, p. 60). Pragmatic rationality is found in the way experts combine not only experiential knowledge and emotional engagement but also scientific evidence and ethical principles for fair processes (Russell, 2017). Expertise or specifically experts who display pragmatic rationality in the rhetorical deliberative setting thus may potentially be held as an alternative guarantor of decision quality.

Concluding, although some authors (Hughes & Doheny, 2011; Jenkings & Barber, 2004) do reference the world outside in the decision-making process, the HSR field primarily focuses its efforts on the dynamics of the deliberative processes and describing the role and embodiment of expertise therein. For this dissertation, these processual dynamics of pragmatic decision making will be of great interest. Based on the overview above, I note three elements: 1) decision-making experts display an understanding of, and are able to work with, different types of knowledges; 2) experts know how to respond in a human way; and 3) experts adhere to formalised procedures. I will,

however, release the emphasis on combining work and instead, focus on the expertise displayed in appraisal specifically and what may distinguish it from earlier studies on coverage decision-making expertise. Moreover, as the use of the necessity criterion is evident not only in deliberation but also in documents containing the justification or rationale for decisions, this study requires more focus in terms of the substance and the outcome of the decision than much current HSR is giving it. To fill this gap, I am turning to Science and Technology Studies.

SCIENCE AND TECHNOLOGY STUDIES: EXPERTISE AND THE WORLD OUTSIDE

Science and Technology Studies (STS) has a long history studying the role of expert advice in 'the greater picture', often a political decision-making process (Collins & Evans, 2008; Fischer, 2011; Frey & Fontana, 1991; King et al., 2018; Rip, 1985). According to Rip (1986, 1992), expert advice should aim to orientate the recipients of the advice in favour of some option for action, involving the world outside with its complexity and uncertainty in this advice (Rip, 1985, p. 95). Rip notes (following Ezrahi, 1980) that when there is no scientific consensus and no agreement on social and political goals, public controversy is likely. The aim should therefore be pragmatic rationalism (see also Rip, 1992), which he defines as:

> Each solution must take account of the circumstances relevant to that unique situation. [Pragmatic rationalism] redefines the goals of expert advice towards stressing its ability to help decision-makers to produce robust outcomes in particular contexts, rather than just in terms of the quality of its scientific and technical content. (Rip, 1985, p. 108)

This definition of pragmatic rationalism differs notably from Russell's. Russell defines pragmatic rationalism as a specific process, a moment in time in which the decision is made, and highlights the skills of experts therein. In contrast, through his definition, Rip underlines the substantive input, the decision outcome, and what happens to the decision afterwards. He defines pragmatic rationalism as a process that is not only localised but in an active relationship with 'contexts'. These contexts affect the decision both through input into the decision as the circumstances to be considered and through testing the 'robustness' of the decision output.

Concerning input into policy decisions, much recent STS work has considered the production of (socially) robust knowledge (Nowotny, 2003; Nowotny et al., 2013; Rip, 2010). Such knowledge, Nowotny et al. hold, is produced by "democratising expertise", which involves opening the door to experts of other kinds, most notably experts-by-experience rather than academic training (cf. Moes, 2019). Context is thus given a voice, "society [is] speaking back to science" (Strathern, 2003), in line with a broader societal trend in many Western countries towards more public participation in policy making (e.g., Jasanoff, 2003). Some scholars have expressed a fear of this resulting in a

collapse of the concept of expertise. Starting with their landmark paper 'The Third Wave of Science Studies', Collins and Evans (2002) have sought to re-draw boundaries around technical-scientific expertise, aiming to explicate its status and role in decision making and sparking subsequent debate within STS (Collins et al., 2010, 2011; Epstein, 2011; Fischer, 2011; Jasanoff, 2003; Rip, 2003).

I would agree with Fischer as he argues that technical expertise inputs in but cannot give the final judgement on decisions affecting the public realm (in line with Lindblom) (Fischer, 2011). Nowotny also gives the many complexities of the social and political world as a major reason for involving experts-by-experience (Nowotny, 2003), thus likewise linking the input to the performance of the output 'out there'. In this, the justification or rationale for a decision likely plays a major role (Bal, 1999). This has also been acknowledged by Collins et al., who noted that political choices based on expert advice ought to be made explicit and public (Collins et al., 2010). As Moreira recommends on coverage decisions, "experts and stakeholders should be able to pre-emptively account for their reasoning to a non-expert audience" (Moreira, 2011, p. 1340). As stated above, the decision justification or rationale should be robust given the places and situations where the decision is to have an effect (Rip, 1985), and knowledge concerning this context is to be crucial in achieving this.

RESEARCH QUESTIONS

This research answers the following main research question: how is the necessity criterion used in practice? The first important observation I made during the early days of my field work, underlined by the policy documents described above, was that the necessity criterion as used in practice took the form of *argumentations*. I will define argumentations as explicated reasons, generally given on paper or in discussion by anyone reasoning about the potential coverage status of a health care technology. Consequently, I set up the following sub-questions:

1. Which primary types of necessity argumentations can be distinguished from the published literature and how does the user and place of use of necessity argumentations affect their use and outcome?

This question, answered in Chapter 1, aims to provide an overview, not comprehensive but to be used as a first insight, of necessity argumentations, to see if a classification can be made that might distinguish several specific argumentation types and if patterns of use may be elucidated based on user and place of use.

2. How are necessity argumentations used in (advised) decision documents in the Netherlands, Belgium, England, and Germany to construct this decision through contextualising the knowledge on a certain health care technology?

Chapter 2 gives the answer to this question, which compares the use of necessity argumentations in (advised) decision documents in Belgium, England, Germany, and the Netherlands to gain insight into how necessity argumentation types are used, for what purpose, and especially how their use compares across four specific case studies.

3. How are necessity argumentations used, both in deliberation and in the advised decision, to 'construct necessity' in the Dutch appraisal committee meetings and what does this say about societal weighing expertise?

For this question, we delved into the setting of the appraisal committee as that is where the use of necessity argumentations such as the ones identified through research question 1 were audible in practice. This question is answered in Chapter 3.

4. How are necessity argumentations used to contribute to a robust advised decision and what may be done to make such a decision more robust?

In Chapter 4, a subset of data from Chapter 3 was re-analysed to formulate a conceptual model for arriving at robust decisions.

Together, these four questions aim to clarify how the necessity criterion is used by examining usage-in-practice as described in scholarly literature, in different countries, in the Dutch appraisal specifically, and from a theoretical stance.

SET-UP AND METHODOLOGY

Chapter 1 comprises a realist review, that answers research question 1. The realist review is an established methodology used to review literatures and to gain insight into technologies-in-context. My team has extended the applicability of this method by applying it to text rather than technology – we are, as far as we know, the first to do so. The realist review method asks: what works, for whom, under what circumstances? This is in fact not difficult to apply to argumentations, as they, like technologies, are used in a particular place to achieve a particular outcome (Pawson et al., 2005). The realist review method was thus chosen to enlarge our understanding of not only the variety of potential argumentations but of their usage, as we hypothesised that not all argumentations would be used in the same way. We chose to review scholarly literature rather than policy documents as much of the previous scholarly work on necessity had in fact built on the latter. Moreover, we were hoping to provide insight into more than the argumentation types used in formalised decisions only. In terms of the review, particularly the identification of the argumentation types was an extended process methodologically. There are iterations of the Excel file that have three argumentation types (Burden of Disease, Necessity of Insurance, and Solidarity, dated 23 March

2016), and there are iterations that have more than the twenty we arrived at (e.g., Vulnerability/Compassion is a combined argumentation type in the final list, as these turned out to be difficult to distinguish in some texts). Chapter 1, then, provides an overview of the list of necessity argumentation types retrieved from scholarly literature, and second, zones in on the patterns of use of these argumentations.

Chapter 2 answers research question 2 by offering a comparative analysis of the use of necessity in decision-making processes generally and in four (advised) decisions specifically across four countries in Western Europe: Belgium, England, Germany, and the Netherlands. These countries were chosen as they share certain 'health care system objectives'; equity, affordability, and transparent decision-making among them (Franken et al., 2012). The reason for examining not just Dutch decision-making practices but comparing them to those in three other countries was that comparative analyses are considered helpful for comparing arguments and statements specifically, resulting in clear definitions and succinct follow-up research questions (Deville et al., 2016). To gain understanding of the general decision-making practices, we started with semi-structured, active group interviews (Holstein & Gubrium, 2016), which were augmented by document and web site analysis, with the results member-checked. This member check also resulted in the decision to additionally examine several decisions that were made in all four countries. Employing the case approach promises insight into patterns of social behaviour specifically (Creswell & Poth, 2017; Ragin, 2004) and is as such a good match for this type of research. The selection of cases adhered to specific criteria formulated by all the collaborators on that chapter (for more details, please see Chapter 2). Importantly, the cases could include any decision as long as a decision document (containing a justification or rationale for the decision) was available, as only this paper was to be examined. The cases that were ultimately selected represented a relatively wide variety of health care technologies, as we hypothesised that would also contribute to the validity of our findings, in addition to the decision outcomes per case varying across the countries. Throughout the analysis of the documentation pertaining to the four selected cases in the four countries, I have used the list of necessity argumentation types as 'sensitising concepts' (Bulmer, 1979) to gain insight into how necessity argumentations are used in coverage decision practice. A sensitising concept is used to make the researcher (more) aware of certain dynamics, and in my case, it guides the analysis through acting as a code. I coded the documents with these twenty necessity argumentation types as codes, transferring them to Microsoft Excel so as to see when, where, and by whom the different argumentation types were used in the documents studied. This has enabled insight not so much in the contents of the argumentations but in the patterns of use to be compared across the four countries.

Chapter 3 will answer research question 3 by homing in on the Dutch appraisal phase and analysing how this appraisal committee interprets its role in terms of societal weighing. Question 3 asks how necessity is constructed, what societal weighing looks like. Chapters 3 and 4, both examine the Dutch appraisal practices specifically, as this is where both the policy documents and Chapter 2 expressly located the use of necessity argumentations. These chapters build in part on

the same cases (namely eculizumab and paracetamol-vitamin D tablets), and case selection was more time-constrained than for Chapter 2. The reason is that I had to have been present at the appraisal committee meeting or at least be able to listen to the audio file, with the former having a strong preference (in the end, only the maternity case happened before my presence at appraisal committee meetings; this case was selected as interviewees stated it would be a particularly fruitful case in terms of studying the necessity criterion). Moreover, the cases were purposely chosen to represent a wide variety of health care technologies. Like in Chapter 2, I coded the transcribed audio files and the documents (in this case, the discussion documents provided as input for the appraisal meeting as well as the decision documents containing the final decision and the letter to the Minister) using the necessity argumentation types formulated in Chapter 1. These data were triangulated by interviews with Institute employees and appraisal committee members.

Sensitisation to the robustness of decisions through appreciation of what public controversy might engender is the way into Chapter 4. This chapter seeks to answer research question 4 to focus more on theorising on how these decisions are brought together in terms of content, asking the question: how are robust coverage decisions made? In this chapter, instead of choosing a primarily data-driven approach, I chose to profit from theoretical work within STS on controversy and robustness to come to a model for making robust decisions in appraisal. This model is illustrated by data from Chapter 3.

REFERENCES

Abrishami, P. (2017). Public Value of Medical Innovations. *A quest for all and for all seasons (PhD-thesis). Maastricht/Diemen: Maastricht University & Zorginstituut Nederland*.

Bærøe, K., & Baltussen, R. (2014). Legitimate Healthcare Limit Setting in a Real-World Setting: Integrating Accountability for Reasonableness and Multi-Criteria Decision Analysis. *Public Health Ethics, 7*(2), 98-111. doi: 10.1093/phe/phu006

Bal, R. A. (1999). Grenzenwerk; over Het Organiseren Van Normstelling Voor De Arbeidsplek.

Baltussen, R., Jansen, M. P., Mikkelsen, E., Tromp, N., Hontelez, J., Bijlmakers, L., & Van Der Wilt, G. J. (2016). Priority Setting for Universal Health Coverage: We Need Evidence-Informed Deliberative Processes, Not Just More Evidence on Cost-Effectiveness. *International Journal of Health Policy and Management, 5*(11), 615-618. doi: 10.15171/ijhpm.2016.83

Baltussen, R., Jansen, M. P. M., Bijlmakers, L., Grutters, J., Kluytmans, A., Reuzel, R. P., . . . Van Der Wilt, G. J. (2017). Value Assessment Frameworks for Hta Agencies: The Organization of Evidence-Informed Deliberative Processes. *Value in Health, 20*(2), 256-260.

Bulmer, M. (1979). Concepts in the Analysis of Qualitative Data. *The Sociological Review, 27*(4), 651-677.

Calnan, M., Hashem, F., & Brown, P. (2017). Still Elegantly Muddling Through? Nice and Uncertainty in Decision Making About the Rationing of Expensive Medicines in England. *International Journal of Health Services, 47*(3), 571-594.

College Voor Zorgverzekeringen. (2006). Pakketbeheer in De Praktijk. Diemen.

College Voor Zorgverzekeringen. (2009). *Pakketbeheer in De Praktijk 2*. Diemen: College voor Zorgverzekeringen.

Collins, H., & Evans, R. (2002). The Third Wave of Science Studies: Studies of Expertise and Experience. *Social Studies of Science, 32*(2), 235-296.

Collins, H., & Evans, R. (2008). *Rethinking Expertise*: University of Chicago Press.

Collins, H., Weinel, M., & Evans, R. (2010). The Politics and Policy of the Third Wave: New Technologies and Society. *Critical policy studies, 4*(2), 185-201.

Collins, H., Weinel, M., & Evans, R. (2011). Object and Shadow: Responses to the Cps Critiques of Collins, Weinel and Evans,''Politics and Policy of the Third Wave'. *Critical policy studies, 5*(3), 340-348.

Commissie-Dunning. (1991). Kiezen En Delen. *Rapport van de Commissie Keuzen in de Zorg*.

Creswell, J. W., & Poth, C. N. (2017). *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. USA: Sage publications.

Dakin, H., Devlin, N., Feng, Y., Rice, N., O'neill, P., & Parkin, D. (2015). The Influence of Cost-Effectiveness and Other Factors on Nice Decisions. *Health Economics, 24*(10), 1256-1271. doi: doi:10.1002/hec.3086

Daniels, N. (2000). Accountability for Reasonableness. *Establishing a fair process for priority setting is easier than agreeing on principles, 321*(7272), 1300-1301. doi: 10.1136/bmj.321.7272.1300

Daniels, N., Porteny, T., & Urrutia, J. (2016). Expanded Hta: Enhancing Fairness and Legitimacy. *International Journal of Health Policy and Management, 5*(1), 1-3. doi: 10.15171/ijhpm.2015.187

Daniels, N., & Sabin, J. (1997). Limits to Health Care: Fair Procedures, Democratic Deliberation, and the Legitimacy Problem for Insurers. *Philosophy & Public Affairs, 26*(4), 303-350. doi: 10.1111/j.1088-4963.1997.tb00082.x

Daniels, N., & Sabin, J. (1998). The Ethics of Accountability in Managed Care Reform. *Health Affairs, 17*(5), 50-64. doi: 10.1377/hlthaff.17.5.50

Daniels, N., & Sabin, J. (2008). Accountability for Reasonableness: An Update. *Bmj, 337*. doi: 10.1136/bmj.a1850

Deville, J., Guggenheim, M., & Hrdličková, Z. (2016). Introduction: The Practices and Infrastructures of Comparison.

Entwistle, V. A., Watt, I. S., Bradbury, R., & Pehl, L. J. (1996). Media Coverage of the Child B Case. *Bmj, 312*(7046), 1587-1591.

Epstein, S. (2011). Misguided Boundary Work in Studies of Expertise: Time to Return to the Evidence. *Critical policy studies, 5*(3), 323-328.

Ezrahi, Y. (1980). Utopian and Pragmatic Rationalism: The Political Context of Scientific Advice. *Minerva, 18*(1), 111-131.

Fischer, F. (2011). The 'Policy Turn'in the Third Wave: Return to the Fact–Value Dichotomy? *Critical policy studies, 5*(3), 311-316.

Franken, M., Le Polain, M., Cleemput, I., & Koopmanschap, M. (2012). Similarities and Differences between Five European Drug Reimbursement Systems. *International Journal of Technology Assessment in Health Care, 28*(4), 349-357. doi: 10.1017/S0266462312000530

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, 119*(2), 195-202. doi: https://doi.org/10.1016/j.healthpol.2014.10.007

Frey, J. H., & Fontana, A. (1991). The Group Interview in Social Research. *The Social Science Journal, 28*(2), 175-187.

Gezondheidsraad. (2003). Contouren Van Het Basispakket. The Hague.

Giacomini, M. K. (1999). The Which-Hunt: Assembling Health Technologies for Assessment and Rationing. *Journal of health politics, policy and law, 24*(4), 715-758.

Gkeredakis, E., Swan, J., Nicolini, D., & Scarbrough, H. (2011). *Rational Judgement Revisited: Practices of Deliberation in Healthcare Funding Decisions*. Paper presented at the OLKC conference, Hull University Business School.

Gordon, E. J. (2006). The Political Contexts of Evidence-Based Medicine: Policymaking for Daily Hemodialysis. *Social Science & Medicine, 62*(11), 2707-2719. doi: https://doi.org/10.1016/j.socscimed.2005.11.024

Guegan, E. W., Huić, M., & Teljeur, C. (2014). Eunethta: Further Steps Towards European Cooperation on Health Technology Assessment. *International Journal of Technology Assessment in Health Care, 30*(5), 475-477.

Guindo, L. A., Wagner, M., Baltussen, R., Rindress, D., Van Til, J., Kind, P., & Goetghebeur, M. M. (2012). From Efficacy to Equity: Literature Review of Decision Criteria for Resource Allocation and Healthcare Decisionmaking. *Cost effectiveness and resource allocation, 10*(1), 9. doi: 10.1186/1478-7547-10-9

Ham, C. (1999). Tragic Choices in Health Care: Lessons from the Child B Case. *Bmj, 319*(7219), 1258-1261. doi: 10.1136/bmj.319.7219.1258

Ham, C., & Glenn, R. (2003). *Reasonable Rationing*: McGraw-Hill Education (UK).

Helderman, J.-K., De Kruijf, J., Verhey, J., & Thiel, S. V. (2014). Dike-Reeve of the Health Care Polder. A Political-Sociological Analysis of the Realisation of the National Health Care Institute against a Backdrop of a Changing Policy Agenda and Changing Political-Dministrative and Societal Relations.

Holstein, J. A., & Gubrium, J. F. (2016). *Narrative Practice and the Active Interview* (Vol. 67): Sage London.

Hughes, D., & Doheny, S. (2011). Deliberating Tarceva: A Case Study of How British Nhs Managers Decide Whether to Purchase a High-Cost Drug in the Shadow of Nice Guidance. *Social Science & Medicine, 73*(10), 1460-1468.

Hughes, D., & Light, D. (2002). *Rationing: Constructed Realities and Professional Practices*. Oxford, UK: Wiley-Blackwell.

Hunter, D. J. (1995). Rationing: The Case for "Muddling through Elegantly". *Bmj, 311*(7008), 811.

Inahta. (2020). What Is Health Technology Assessment (Hta)?

Jansen, M. P., Helderman, J.-K., Boer, B., & Baltussen, R. (2017). Fair Processes for Priority Setting: Putting Theory into Practice : Comment on "Expanded Hta: Enhancing Fairness and Legitimacy". *International Journal of Health Policy and Management, 6*(1), 43-47. doi: 10.15171/ijhpm.2016.85

Jasanoff, S. (2003). Breaking the Waves in Science Studies: Comment on Hm Collins and Robert Evans,the Third Wave of Science Studies'. *Social Studies of Science, 33*(3), 389-400.

Jenkings, K. N., & Barber, N. (2004). What Constitutes Evidence in Hospital New Drug Decision Making? *Social Science & Medicine, 58*(9), 1757-1766.

Kalf, R. R., Makady, A., Ten Ham, R. M., Meijboom, K., Goettsch, W. G., & Workpackage, O. B. O. I.-G. (2018). Use of Social Media in the Assessment of Relative Effectiveness: Explorative Review with Examples from Oncology. *JMIR cancer, 4*(1), e11.

Kapiriri, L., Norheim, O. F., & Martin, D. K. (2009). Fairness and Accountability for Reasonableness. Do the Views of Priority Setting Decision Makers Differ across Health Systems and Levels of Decision Making? *Social Science & Medicine, 68*(4), 766-773. doi: https://doi.org/10.1016/j.socscimed.2008.11.011

Kapiriri, L., & Razavi, D. (2017). How Have Systematic Priority Setting Approaches Influenced Policy Making? A Synthesis of the Current Literature. *Health Policy, 121*(9), 937-946. doi: https://doi.org/10.1016/j.healthpol.2017.07.003

King, N., Horrocks, C., & Brooks, J. (2018). *Interviews in Qualitative Research*. UK: SAGE Publications Limited.

Lehoux, P. (2014). *The Problem of Health Technology*. USA: Routledge.

Lehoux, P., & Blume, S. (2000). Technology Assessment and the Sociopolitics of Health Technologies. *Journal of health politics, policy and law, 25*(6), 1083-1120.

Lindblom, C. E. (1959). The Science of Muddling Through. *PublicAdministrationReview, 19*(2), 79-88.

Makady, A., Ten Ham, R., De Boer, A., Hillege, H., Klungel, O., & Goettsch, W. (2017). Policies for Use of Real-World Data in Health Technology Assessment (Hta): A Comparative Study of Six Hta Agencies. *Value in Health, 20*(4), 520-532.

Martin, D., Giacomini, M., & Singer, P. (2002). Fairness, Accountability for Reasonableness, and the Views of Priority Setting Decision-Makers. *Health Policy, 61*(3), 279-290.

Martin, D., Pater, J., & Singer, P. (2001). Priority-Setting Decisions for New Cancer Drugs: A Qualitative Case Study. *The Lancet, 358*(9294), 1676-1681. doi: https://doi.org/10.1016/S0140-6736(01)06714-9

Mechanic, D. (1997). Muddling through Elegantly: Finding the Proper Balance in Rationing. *Health Affairs, 16*(5), 83-92.

Moes, F. (2019). *Experts and Idiots: An Ethnography of Epistemic Practices in Dutch Social Health Insurance*: Maastricht University.

Moes, F., Houwaart, E., Delnoij, D., & Horstman, K. (2016). Contested Evidence: A Dutch Reimbursement Decision Taken to Court. *Health Economics, Policy and Law, 12*(3), 325-344. doi: 10.1017/S1744133116000281

Moreira, T. (2005). Diversity in Clinical Guidelines: The Role of Repertoires of Evaluation. *Social Science & Medicine, 60*(9), 1975-1985.

Moreira, T. (2011). Health Care Rationing in an Age of Uncertainty: A Conceptual Model. *Social Science & Medicine, 72*(8), 1333-1341.

Morrell, L., Wordsworth, S., Fu, H., Rees, S., & Barker, R. (2017). Cancer Drug Funding Decisions in Scotland: Impact of New End-of-Life, Orphan and Ultra-Orphan Processes. *BMC health services research, 17*(1), 613.

Niëns, L. (2014). Affordability in Health Care: Operationalizations and Applications in Different Contexts.

Niezen, M. M. (2012). *Working Towards Legitimacy in Decision Making: On Ogoverning Appropriate Medicine Use and Reimbursement in Health Care*.

Nowotny, H. (2003). Democratising Expertise and Socially Robust Knowledge. *Science and public policy, 30*(3), 151-156.

Nowotny, H., Scott, P. B., & Gibbons, M. T. (2013). *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty*. Cambridge, UK: John Wiley & Sons.

Oliver, A., Mossialos, E., & Robinson, R. (2004). Health Technology Assessment and Its Influence on Health-Care Priority Setting. *International Journal of Technology Assessment in Health Care, 20*(1), 1-10. doi: 10.1017/S026646230400073X

Oortwijn, W., Jansen, M., & Baltussen, R. (2020). Use of Evidence-Informed Deliberative Processes by Health Technology Assessment Agencies around

the Globe. *International Journal of Health Policy and Management, 9*(1), 27.

Patera, N., & Wild, C. (2014). Assessment – Appraisal – Decision. Lbi-Hta Decision Support Document Nr.: 72. Wien: Ludwig Boltzmann Institut für Health Technology Assessment.

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist Review–a New Method of Systematic Review Designed for Complex Policy Interventions. *Journal of Health Services Research & Policy, 10*(suppl 1), 21-34.

Poley, M. J., Stolk, E.A., Brouwer, W. B. F., Van Busschbach,J. J. . (2002). Ziektelast Als Uitwerking Van Het Criterium 'Noodzakelijkheid' Bij Het Makenvan Keuzen in De Zorg. *Ned Tijdschr Geneeskd, 146*(4), 2312-2315.

Raad Voor De Volksgezondheid En Zorg. (2006). Zinnige En Duurzame Zorg. Zoetermeer.

Raad Voor De Volksgezondheid En Zorg. (2007). Rechtvaardige En Duurzame Zorg. Zoetermeer.

Ragin, C. C. (2004). Turning the Tables: How Case-Oriented Research Challenges. *Rethinking social inquiry: Diverse tools, shared standards, 123*.

Reckers-Droog, V., Van Exel, J., & Brouwer, W. (2018). Looking Back and Moving Forward. *Health Policy*.

Reckers-Droog, V., Van Exel, J., & Brouwer, W. (2019). Equity Weights for Priority Setting in Healthcare: Severity, Age, or Both? *Value in Health, 22*(12), 1441-1449.

Rip, A. (1985). Experts in Public Arenas. *Regulating industrial risks: Science, hazards and public protection*, 94-110.

Rip, A. (1986). Controversies as Informal Technology a Ssessment. *Knowledge, 8*(2), 349-371.

Rip, A. (1992). Expert Advice and Pragmatic Rationality. *The Culture and power of knowledge: inquiries into contemporary societies*, 363.

Rip, A. (2003). Constructing Expertise: In a Third Wave of Science Studies? *Social Studies of Science, 33*(3), 419-434.

Rip, A. (2010). Social Robustness and the Mode 2 Diagnosis. *Science, Technology & Innovation Studies, 6*(1). doi: http://dx.doi.org/10.17877/DE290R-13458

Russell, J. (2017). *The Rationality of Rationing: A Rhetorical Policy Analysis of Deliberations About Resource Allocation in the Nhs.* University of Oxford, UK.

Russell, J., & Greenhalgh, T. (2014). Being 'Rational'and Being 'Human': How National Health Service Rationing Decisions Are Constructed as Rational by Resource Allocation Panels. *Health:, 18*(5), 441-457.

Russell, J., Swinglehurst, D., & Greenhalgh, T. (2014). 'Cosmetic Boob Jobs' or Evidence-Based Breast Surgery: An Interpretive Policy Analysis of the Rationing of 'Low Value'treatments in the English National Health Service. *BMC health services research, 14*(1), 413.

Shah, K. K. (2009). Severity of Illness and Priority Setting in Healthcare: A Review of the Literature. *Health Policy, 93*(2-3), 77-84.

Singer, P. A., Martin, D., Giacomini Bhatia, M. V., & Purdy, L. (2000). Priority Setting for New Technologies in Medicine: Qualitative Case Study. *Bmj, 321*(7272), 1316-1318. doi: 10.1136/bmj.321.7272.1316

Stolk, E., Brouwer, W., & Busschbach, J. (2002). Rationalising Rationing: Economic and Other Considerations in the Debate About Funding of Viagra. *Health Policy, 59*(1), 53-63.

Stolk, E., De Bont, A., Van Halteren, A., Bijlmer, R., & Poley, M. (2009). Role of Health Technology Assessment in Shaping the Benefits Package in the Netherlands. *Expert review of pharmacoeconomics & outcomes research, 9*(1), 85-94.

Stolk, E., Van Donselaar, G., Brouwer, W., & Busschbach, J. (2004). Reconciliation of Economic Concerns and Health Policy. *PharmacoEconomics, 22*(17), 1097-1107.

Strathern, M. (2003). Re-Describing Society. *Minerva, 41*(3), 263-276.

Vuorenkoski, L., Toiviainen, H., & Hemminki, E. (2008). Decision-Making in Priority Setting for Medicines—a Review of Empirical Studies. *Health Policy, 86*(1), 1-9.

Walley, T. (2007). Health Technology Assessment in England: Assessment and Appraisal. *Medical Journal of Australia, 187*(5), 283.

Wiering, B., De Boer, D., & Delnoij, D. (2017). Patient Involvement in the Development of Patient-Reported Outcome Measures: A Scoping Review. *Health Expectations, 20*(1), 11-23.

Zorginstituut Nederland. (2013). Pakketbeheer in De Praktijk Deel 3 (pp. 1-106). Diemen.

Zorginstituut Nederland. (2016). Reglement Adviescommissie Pakket. In Z. Nederland (Ed.). https://www.zorginstituutnederland.nl/over-ons/publicaties/besluit/2016/12/05/reglement-adviescommissie-pakket-zorginstituut-nederland (accessed 11 February 2020).

Zorginstituut Nederland. (2017). Pakketadvies in De Praktijk: Wikken En Wegen Voor Een Rechtvaardig Pakket. Diemen.

Zorginstituut Nederland. (2018). Ziektelast in De Praktijk. Diemen.

Zorginstituut Nederland. (2020). Academische Werkplaats Hta&A. Retrieved 28 January, 2020

1 Realist review of necessity argumentations

Published as: *Kleinhout-Vliek, T., de Bont, A., & Boer, B. (2017). The bare necessities? A realist review of necessity argumentations used in health care coverage decisions. Health Policy, 121(7), 731-744.*

INTRODUCTION

Public outrage often ensues when decision makers exclude forms of care, such as orphan drugs or expensive cancer medicines, based on an incremental cost-effectiveness ratio (ICER) that is below par. This outrage not infrequently precedes a reversal of the decision (Burls et al., 2005; Clarke et al., 2001). At the same time, however, not all forms of care with a sufficiently low ICER are covered. Viagra, for example, is highly effective and not that expensive but almost never provided by the state; decision makers deem it unnecessary to do so (Bernfort, 2003; Stolk et al., 2002). Hence, (cost-) effectiveness is not the decisive factor in all funding decisions. In these situations, another factor trumps it: the perceived *necessity* of coverage. To aid operationalisation, this chapter will survey the content, use, and context of the necessity criterion, an umbrella term for need- and solidarity-related argumentations used – not just decisively, and not just in coverage decisions made by policy makers and insurance companies.

In this chapter we will review argumentations underpinning the necessity, or lack thereof, of coverage of a certain treatment or therapy, as explicated in academic literature. To cast our net wide, we have chosen to include not only 'actual' decisions, that is, coverage decisions made by policy makers and insurance companies but also what we term 'hypothetical' coverage decisions. The latter type generally comes in the form of surveys (of, e.g., decision makers or the public) or ethical or economic analyses exploring possible reasons for (denial of) coverage. By examining both actual and hypothetical decisions we hope to provide insight into all potential considerations that may be invoked when deciding whether the coverage of a therapy or treatment is thought to be necessary. This is relevant as surveys and public opinion are considered of note (and of use) within coverage decision making practice (Ham, 1997; Mitton et al., 2009), as are scholarly reflections, as exemplified by the international take-up of the Accountability for Reasonableness framework (Kapiriri et al., 2009; D. Martin et al., 2002).

Objectives and focus of review

We followed the realist review method as described in the RAMESES publication standard (G. Wong et al., 2013). This method is used to review sundry literatures on a specific policy intervention, in order to describe why and how these interventions do what they do in their context. Using this method, researchers aim to uncover what works, for whom, and in what circumstances by conceptualising meta-level theories that detail patterns of how mechanisms-in-contexts lead to certain outcomes (Greenhalgh et al., 2011; Otte-Trojel et al., 2014; Pawson et al., 2005). In such an iterative research process, we refined how argumentations bearing upon necessity of coverage (mechanism) are used in justifying both actual and hypothetical coverage decisions (outcome), as found in academic literature. These decisions are made in context: by different decision makers from different countries, and, in case of actual decisions, placed on the decision agenda by different actors. Thus, the argumentations may be seen as interventions that have a proposed or actual outworking, also depending on contexts they are situated in. This review will address the following questions:

1. Which, if any, argumentations (mechanisms) are currently used in hypothetical and actual coverage decisions to justify whether coverage of a treatment is, or is not, necessary?
2. How do these argumentations justify the hypothetical and actual coverage decisions (outcomes) for different treatments, in different countries, put on the agenda by different agents (contexts)?

METHODS

Rationale for using realist synthesis

Little attention has been given to "the problem of operationalizing for decision makers essentially qualitative and normative criteria such as whether the technology serves an "ethical" or "medically necessary" purpose" (Giacomini, 2005). Furthermore, "social and ethical parameters of value (…) are anticipated to become as critical for reimbursement decisions (…) as economic and clinical criteria" (Akhmetov & Bubnov, 2015). In light of the lack of operationalisation and its (potentially) crucial role in coverage decisions, we conducted a literature review of the argumentation types that fall under the necessity criterion.

A realist review describes an intervention from different types of literatures, in our case actual coverage decisions (qualitative analyses of coverage decisions or policies), as well as hypothetical ones (economic analyses, ethical analyses, surveys, interviews, and opinion pieces). It searches these articles not just for information on the intervention (that is, the argumentation) but also for how the context (country, agenda setter) may have influenced the use of the intervention and its outcome (the decision including decision type: hypothetical or actual decision). This is subsequently summarised in context-mechanism-outcome patterns. From these patterns, meta-level theories are formulated that explain the working of these interventions-in-context. The primary reason for choosing the realist review method is practical; this method provided a focused lens to zoom in on particular aspects of actual and hypothetical coverage decisions, which in turn aided comparison of a broad variety of articles. Using this method for a non-classical intervention proved, moreover, an interesting methodological issue to grapple with. The second reason for utilising this method lies in its philosophical underpinnings. A realist philosophy holds that actors can and do effectuate change in context but are themselves shaped by the contexts they are part of. In this sense, it is likely to be acceptable to (social) scientists and policy makers alike.

Scoping the literature and searching processes

As an exploratory foray into grey literature and policy documents yielded too few explicated argumentations, we focused on peer-reviewed literature. For our primary background search thereof (Pawson et al., 2005), we used the conceptualisation of the necessity criterion in the Netherlands (Couwenbergh et al., 2013), as a request for operationalisation of this criterion from the Dutch National Health Care Institute catalysed this study. We subsequently discovered similar and/or underlying conceptualisations and related terminology in other countries, like 'need' and 'solidarity', which helped inform our search

strategy. The primary search was conducted in Embase (see Table 1 for search terms) and translated to Medline and Web of Science (Bramer et al., 2014), which is recognised to be an effective combination for reviews (Wichor M. Bramer). We used three general elements separated by the Boolean operator 'AND' as this kept the total number of articles workable (under 6,000). These general elements are a) the type of provision, b) the process of decision making, and c) the content in terms of criteria. Utilising a), we aimed for a representative sample, therefore a wide variety of provision were included (benefit package, health insurance, and/or health catalogue or service). For elements b) and c) specificity was the goal; we zoomed in specifically on coverage decisions (also often termed 'rationing' or 'priority setting' decisions) and precisely on those decisions that employ the necessity criterion. In selecting the exact search terms, we aimed for results that included the articles retrieved and selected from the primary background search, for example (Bernfort, 2003; Hoedemaekers & Oortwijn, 2003; Stolk et al., 2002). For each of the three elements we included relevant thesaurus terms (Emtree terms for Embase and MeSH terms for Medline). We excluded conference papers, letters, notes, and editorials, as well as articles written in any language other than English but did not employ any date restrictions (WM Bramer).

('insurance'/de OR 'health insurance'/de OR 'child health insurance'/de OR 'national health insurance'/de OR 'private health insurance'/de OR 'public health insurance'/de OR 'national health service'/de OR 'reimbursement'/de OR (insurance* OR reimburse* OR (national NEAR/3 (service OR coverage)) OR (cover* NEAR/6 deci*) OR ((partial* OR polic* OR universal OR unlimited OR limited OR temporar* OR permanent* OR recommend* OR plan OR plans) NEAR/3 coverage) OR (basic NEAR/3 package*) OR (health NEAR/3 catalogue*)):ab,ti)
AND
('decision making'/de OR 'ethical decision making'/de OR 'medical decision making'/de OR (decision* OR decide OR rationing OR priorit* OR (analys* NEAR/3 (inclusion OR exclusion)) OR (coverage NEAR/3 (negativ* OR positiv* OR determin* OR deny OR denial*))):ab,ti) AND ('resource allocation'/de OR (coverage OR inclusion* OR funding OR (resource* NEAR/3 allocat*) OR 'should be provided' OR 'what to provide'):ab,ti)
AND
('health care cost'/de OR 'cost of illness'/de OR 'economic evaluation'/exp OR ethics/de OR bioethics/de OR 'medical ethics'/de OR 'ethical decision making'/de OR 'health care policy'/de OR 'needs assessment'/de OR (necess* OR cost* OR (disease* NEAR/3 burden*) OR expenditure* OR solidarit* OR (therapeutic NEAR/3 (value* OR need*)) OR (budget* NEAR/3 impact*) OR ethic* OR 'health benefit*' OR (benefit NEAR/3 (risk OR analysis)) OR 'health technology assessment*' OR 'health care poli*' OR (need* NEAR/3 (assess* OR healthcare OR health-care))):ab,ti)
NOT
([Conference Abstract]/lim OR [Letter]/lim OR [Note]/lim OR [Editorial]/lim)
AND
[english]/lim

Table 1 - Search terms used in Embase

Selection and appraisal of documents

The first author (TKV) scanned titles, abstracts, and keywords in Endnote to include decisions that were made on the macro (government) or meso (local health authorities, sickness funds, and insurance companies) level (Polikowski & Santos-Eggimann, 2002). BB, the last author, scanned a random subset of 537 studies. Together, an agreement rate of 96% was reached and further disagreement was resolved through discussion. This first round of inclusions amounted to a total of 666 studies. Next, TKV read all candidate papers in full and excluded 594 of the 666, ending up with 72 studies. Through snowballing, a further 26 such studies were added, bringing the total number to 98 (see Figure 1).

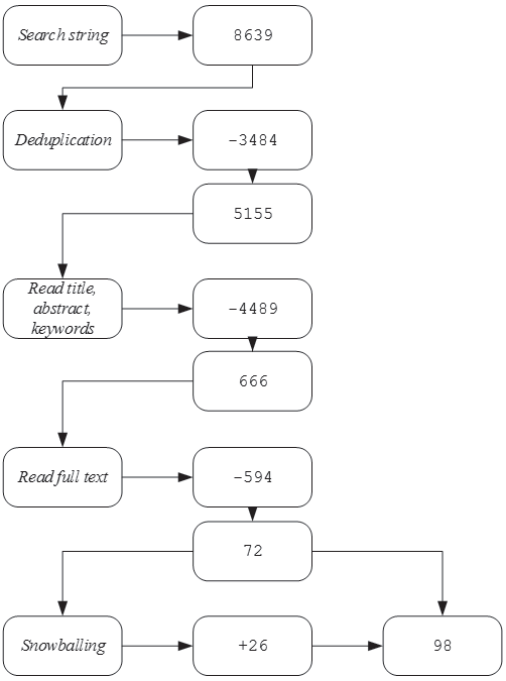

Figure 1 - Document flow diagram

The inclusion criteria were:

a) the article describes a hypothetical or actual coverage decision made concerning a specific form of care, treatment, therapy or medical technology,
b) on the meso or macro level, and
c) employing necessity-related considerations.

We correspondingly excluded articles that:

a) describe generalised criteria not applied to a specific form of care,
b) concern individual decisions (that is, for one patient) as they may concern the exception rather than the rule, and
c) employ only effectiveness and/or cost-effectiveness or other quantitative argumentations.

In doing so, we aimed to find studies containing concrete, qualitative, in-use argumentations that influence different types of coverage decisions as formulated by different actors on different levels regarding different forms of care.

Data extraction

TKV extracted data from these 98 articles using Microsoft Excel and ATLAS.ti. During this process, 'mode of agenda setting' emerged as a potential context of influence. Conversely, 'type of disease'

was removed as it proved unfruitful. To score the context, we noted country, year, macro- or meso level, and decision maker per identified argumentation, which was in turn scored as mechanism. The coverage decision was scored as the outcome, with eight outcome categories, of which three were used in case of an actual coverage decision:

- Yes – the treatment is unconditionally covered/conditions unspecified,
- Conditional – the treatment is covered for specific patient groups/indications,
- No – the treatment is never covered.

Three categories were scored as hypothetical decisions (generally based on theory, interviews, or surveys):

- Should – the treatment should be unconditionally covered/conditions unspecified,
- Should conditional – the treatment should be covered for specific patient groups or indications,
- Should not – the treatment should never be covered.

Furthermore, two categories were added later on to signify when an author or actor noted whether an argumentation should, or should not, have a bearing upon the decision.

- Valid – the consideration should be taken into account,
- Not valid – the consideration should not be taken into account.

BB extracted data from a random subset of 11 studies, reaching an agreement rate of 73%, and full agreement was reached in deliberation.

Analysis and synthesis process

Data analysis was undertaken by TKV, and candidate context-mechanism-outcome combinations were formulated according to the following pattern:

Based on these argumentation types,	(mechanism)
used in country C, put on the agenda by A,	(contexts)
lead to actual or hypothetical in- or exclusion of treatment T.	(outcome)

Discussion with BB and AdB enabled the formulation of four meta-level theories.

RESULTS

Document characteristics

The 98 reviewed studies contain studies that are (in part) based on qualitative analyses of coverage decisions or policies (34 studies), interviews/surveys (26 studies), and case studies (twenty studies), but also ethical analyses (fourteen studies), opinion pieces (fourteen studies), economic analyses (eight studies), and reviews (seven studies), from over twenty (primarily Western) countries.

Main findings

General overview

In this chapter twenty argumentation types are described that are in use in hypothetical and actual coverage decisions. These twenty argumentation types are derived from over 400 argumentations found in 98 articles, for an overview of the argumentation types in alphabetical order, see Table 2.

Argumentation type	Description
1) Definition of Illness	Whether the ailment is considered an illness for which treatment is necessary
2) Dignity	Whether (lack of) coverage is considered to affect the dignity of the patient to such an extent that it needs to be amended
3) Equity/Fairness/Justice	Whether coverage would be necessary to counter injustice/inequity/lack of fairness in (access to) treatment
4) Human Right	Whether (lack of) coverage is considered to affect the human rights of the patient to such an extent that it needs to be amended
5) Individual Cost	Whether lack of coverage would stop patients from buying necessary care themselves due to prohibitive cost
6) Individual Responsibility	Whether the individual is considered responsible for paying for this treatment
7) Medical Necessity	Whether or not a treatment is considered to be "medically necessary" or a "medical necessity"
8) Morbidity/Severity	Whether the physical and/or psychosocial morbidity associated with a certain ailment constitutes such a need that coverage is considered necessary
9) Moral Hazard	Whether there is considered to be a possibility of over-usage (i.e., unnecessary increase in demand, when people use more than they need as a result of coverage)
10) Need	The extent to which the patient is considered to be in need for which treatment is necessary
11) (No) Alternative	Whether or not viable alternatives are considered to be present which would make coverage more or less necessary
12) Patient-Diagnosis	Whether an illness is self-reported rather than diagnosed by a doctor
13) Range of Normality	Whether the experience of the patient is considered normal or abnormal to such an extent that coverage is deemed necessary
14) Rule of Rescue	Whether the identifiability of individuals close to death is considered to heighten the necessity of coverage
15) Similar Treatments	Whether similar treatments are covered or not (meaning that this *type* of treatment is considered necessary)
16) Small Number of Patients	Whether the small size of the patient population is considered to heighten the necessity of coverage (due to, amongst others, the inequality in terms of research expenditure or difficulties in obtaining high-quality data)
17) Societal Impact	Whether coverage is considered necessary to allay the impact this disease has on people beyond the patient
18) Societal Functioning	Whether coverage would aid a person's necessary functioning in society
19) Societal Responsibility	Whether society is considered responsible for this necessary treatment
20) Vulnerability/Compassion	Whether a compassionate response to vulnerable groups, e.g. children, in the form of coverage is considered to be a necessity

Table 2 - Overview of the twenty argumentation types that fall under the necessity criterion and their respective descriptions

The 98 reviewed articles are divided up into seven treatment sets (cancer therapies, orphan drugs, infertility treatments, Viagra, cosmetic surgery, obesity treatment, and smoking cessation therapy). What follows is a detailed description of the use of these twenty argumentation types per treatment set in justifying positive, negative, and/or conditional coverage decisions, both actual and hypothetical. A handful of included articles is not described because there was too little information on that type of treatment: Alzheimer's disease (Bernardi & Pegoraro, 2003), genetic tests (Fischer & Rogowski, 2014; Trosman et al., 2015), and medical devices (Kisser et al., 2016). Notably, no new argumentation types are described in these articles.

1. Cancer therapies

For cancer therapies, we observed solely positive coverage decisions. For these, higher incremental cost-effectiveness ratios than usual are acceptable (Chabot & Rocchi, 2010), as evidenced by positive decisions made despite the fact that coverage would be inadvisable based on high cost (Drummond & Mason, 2007), or "lack of clear clinical consensus about [a therapy's] benefits" (Daniels & Sabin, 1998). Moreover, a few articles pose that "patients should be treated equally" regardless of the number of patients suffering from the specific type of cancer.

The main argumentation visible in this data subset poses that the need of the patient and the severity of the disease should be taken into account (Cookson et al., 2008; Linley & Hughes, 2013; Rocchi et al., 2008; Stafinski et al., 2014; Trowman et al., 2011). As cancer patients have urgent needs and a serious health condition, therefore therapy is, or should be, covered (Aggarwal et al., 2014; Brock, 2010; Cookson et al., 2008; Daniels & Sabin, 1998; Drummond et al., 2009; Drummond & Mason, 2007; D. Martin et al., 2001; Rocchi et al., 2008; Singer et al., 2000), heightened by concerns over dignity and impact on daily activities and family (Stafinski et al., 2014). The Rule of Rescue is the second main argumentation, where the necessity of covering a certain treatment is high specifically for identifiable patients who are in a life-threatening situation and are without alternative (Brock, 2010; MacKenzie et al., 2008; Singer et al., 2000). The unavailability of alternative treatments or therapies should thus be considered (Cookson et al., 2008; Daniels & Sabin, 1998; Drummond et al., 2009; Lim et al., 2014; Linley & Hughes, 2013), as visible in several coverage decisions (Chabot & Rocchi, 2010; Cookson et al., 2008; Daniels & Sabin, 1998; Drummond et al., 2009; Drummond & Mason, 2007; Lim et al., 2014; Linley & Hughes, 2013; D. Martin et al., 2001; Singer et al., 2000), but this argumentation is also sometimes questioned (Brock, 2010; Cookson et al., 2008). The 'identifiable individuals' element is also present in the 'Small Number of Patients' argumentation, which several authors pose is important (Cookson et al., 2008; Stafinski et al., 2014; Trowman et al., 2011) and as heightening justifiability (Chabot & Rocchi, 2010; Drummond et al., 2009), although others deny this (Cookson et al., 2008; Linley & Hughes, 2013; Singer et al., 2000). The third main argumentation holds that for some, like children, exceptions are (Lim et al., 2014; Rocchi et al., 2008), may (Cookson et al., 2008; Stafinski et al., 2014), may not (Linley & Hughes, 2013), or should be made (Brock, 2010). This is expanded by Brock, who argues that those who have had fewest QALYs in their lifetime are the worst off, which is most obviously the case

for paediatric patients (Brock, 2010). The final argumentation concerns the distribution of health care resources (Aggarwal et al., 2014; Drummond & Mason, 2007). E.g., Foy et al. emphasise that patients in different districts should be treated equally, thus arguing for more distributive justice (Foy et al., 1999).

2. Orphan drugs

For orphan drugs, we identified argumentations for both positive and negative coverage decisions. The structure and types of argumentations of positive decisions in this data subset greatly resemble those for cancer therapies and the use of a number of argumentations is questioned here as well.

Higher incremental cost-effectiveness ratios are generally found acceptable for orphan drugs so decisions reek of making exceptions to this rule. The fact that the severity of the illness does or should count as an argumentation is stressed by many (Bae et al., 2015; Clarke et al., 2001; Cohen & Felix, 2014; Denis et al., 2011; Dyfrig Hughes, 2006; Mentzakis et al., 2011; Nicod, 2016; Rosenberg-Yunger et al., 2011; Schlander & Beck, 2009; Winquist et al., 2012; Zelei et al., 2016), and so is need (Cohen & Felix, 2014; Nicod, 2016). Consequently, argumentations in favour of coverage include that orphan diseases are severe (Drummond & Towse, 2014; Dyfrig Hughes, 2006; Largent & Pearson, 2012; Paulden et al., 2015), that coverage only happens when they are (Clarke et al., 2001) and/or when patients are in need (Henschke, 2012). Patients are said to be vulnerable, which should be (Largent & Pearson, 2012; Winquist et al., 2012), and is (Bañón Hernández Antonio, 2015; Henschke, 2012), considered, and to score low on societal functioning, and orphan drugs should be covered because they maintain or restore these capacities (Largent & Pearson, 2012). As for cancer therapies, a common argumentation supportive of provision is that patients have "no alternative" (Drummond & Towse, 2014; Iskrov et al., 2013; Paulden et al., 2015), often generally considered important (Cohen & Felix, 2014; Denis et al., 2011; Rosenberg-Yunger et al., 2011; Zelei et al., 2016). The Rule of Rescue argues in favour of coverage in two cases (Bañón Hernández Antonio, 2015; Clarke et al., 2001), its relevance as a consideration is underlined by some (Bae et al., 2015; Rosenberg-Yunger et al., 2011; Zelei et al., 2016) but also regularly questioned (Burls et al., 2005; Arna S Desser et al., 2010; Gross, 2002; David Hughes et al., 2005; Juth, 2014; Largent & Pearson, 2012; Schlander & Beck, 2009). The small patient population is mentioned as an argumentation separately as well but primarily to say it should not be of effect on the coverage decision (Burls et al., 2005; Arna S. Desser, 2013; Arna S Desser et al., 2010; Drummond & Towse, 2014; David Hughes et al., 2005; Juth, 2014; McCabe et al., 2005; Mentzakis et al., 2011; Paulden et al., 2015), though not always (Dyfrig Hughes, 2006; Nicod, 2016; Zelei et al., 2016). Finally, the societal impact of treatment should be considered (Zelei et al., 2016). A new main argumentation in favour of coverage holds that the budget impact remains "sufficiently insignificant" given the low numbers of patients (David Hughes et al., 2005), as is the concurrent fact that the price is prohibitive to the individual (Denis et al., 2011). "Equity" and distributive justice are once again mentioned but the exact line of reasoning is not always explicated (Burls et al., 2005; David Hughes et al., 2005; Rosenberg-Yunger et al., 2011; Zelei et al., 2016), as is the case for the right to health care (David Hughes et al., 2005).

Arguing for denial of coverage, the most common assertion is that some orphan diseases may not be as severe as assumed, e.g. Gaucher's disease is said to be "minimally symptomatic" (Clarke et al., 2001; Gross, 2002). Hughes et al. state that orphan diseases may not "pose *sufficient imminent threat* to the life of patients to constitute a right to treatment" (italics mine) (David Hughes et al., 2005), see also (Juth, 2014). One coverage decision cites "equity" as a reason (Rosenberg-Yunger et al., 2011). Finally, denial of coverage for similar treatments may support another negative coverage decision (Burls et al., 2005; Winquist et al., 2012).

3. Infertility treatments

The infertility treatment data subset was the largest and the most diverse, containing positive coverage decisions, including those that set indication criteria, as well as negative coverage decisions.

The first main argumentation favouring coverage is that those seeking infertility treatment are in need (Blank, 1997; McMillan, 2001; McWhirter & McQueen, 2000; Rauprich et al., 2010; Redmayne & Klein, 1993; Shaw et al., 2002), resulting in a positive coverage decision once (Redmayne & Klein, 1993). Further, they are suffering morbidity (Brown, 2000; Chambers et al., 2013; E. G. Hughes & Giacomini, 2001; Lord et al., 2001; Redmayne & Klein, 1993; Shaw et al., 2002), explicated as mental distress and psychological harm in decisions by UK health authorities (Redmayne & Klein, 1993). Treatment may be medically necessary (Nachtigall et al., 2012) and patients therefore deserve compassion (Giacomini et al., 2000; Nachtigall et al., 2012). The second argumentation holds that the "cost of [coverage of] assisted conception would be surprisingly small to the NHS as a whole, although to an individual it is often prohibitive" (Lord et al., 2001) and (Chambers et al., 2013; Chambers et al., 2006; E. G. Hughes & Giacomini, 2001). Third, equity is mentioned again, but the underlying line of reasoning is more clearly explicated here, though not used in actual decisions. Lack of geographical equity as well as equity over the rich-poor divide is cause for concern (Brown, 2000; Chambers et al., 2006; Giacomini et al., 2000; E. G. Hughes & Giacomini, 2001; Lord et al., 2001; McWhirter & McQueen, 2000; Nachtigall et al., 2012; Redmayne & Klein, 1993; Shaw et al., 2002). Postcode and financial status should not affect coverage, and it would therefore be preferable to cover infertility treatment everywhere. Moreover, it is deemed inequitable to withhold treatment based on the environmental impact of having a child (Wilkinson & Williams, 2015). The fourth and fifth argumentations are not visible in the first two datasets, and both are only in use in hypothetical decisions, made by authors, patients, and the public. The fourth holds that infertility, or subfertility (E. G. Hughes & Giacomini, 2001), is an illness and its cure should therefore be covered (Brown, 2000; Chambers et al., 2013; Giacomini et al., 2000; E. G. Hughes & Giacomini, 2001; Lord et al., 2001; McMillan, 2001; McWhirter & McQueen, 2000; Mladovsky & Sorenson, 2010; Nachtigall et al., 2012; Neumann, 1997; Plomer et al., 1999; Rauprich et al., 2010; Redmayne & Klein, 1993; Wilkinson & Williams, 2015), especially because similar services are covered (E. G. Hughes & Giacomini, 2001; Lord et al., 2001; Wilkinson & Williams, 2015). Fifth, parenthood is part of the "right to reproduce" (Giacomini et al., 2000) or right to health (Mladovsky & Sorenson, 2010), required for societal functioning (Chambers et al., 2013; Lord et al., 2001; McMillan, 2001;

Mladovsky & Sorenson, 2010; Neumann, 1997; Redmayne & Klein, 1993) and part of "the basic opportunities every human should have" (Rauprich et al., 2010), also (Blank, 1997) and that those without children suffer stigmatisation (Chambers et al., 2013; Lord et al., 2001) and their dignity is impacted (E. G. Hughes & Giacomini, 2001). Lastly, the societal impact of children, in economic terms, should be considered (Connolly et al., 2010).

For the negative coverage decisions, we observe several argumentations, of which some are new. First, infertility is not an illness and therefore not covered (Neumann, 1997; Redmayne & Klein, 1993) or should not be (Wilkinson & Williams, 2015), also because it is not life-threatening (Neumann, 1997). In some cases, it is only covered in case of a diagnosed cause (Brown, 2000; McWhirter & McQueen, 2000). Other needs are thought "more pressing" (Giacomini, 2005; E. G. Hughes & Giacomini, 2001), the treatment risky and not medically necessary (E. G. Hughes & Giacomini, 2001), also the reason for exclusion from coverage in Ontario (Giacomini et al., 2000). Infertility has little emotional appeal (E. G. Hughes & Giacomini, 2001) and its treatment should therefore fall to personal, and not societal, responsibility (Blank, 1997; E. G. Hughes & Giacomini, 2001; Plomer et al., 1999). Other argumentations that support a hypothetical negative coverage decision include potential moral hazard (Blank, 1997; Mladovsky & Sorenson, 2010; Neumann, 1997) and that alternatives are available (Nachtigall et al., 2012; Wilkinson & Williams, 2015), of which the latter also used in an actual decision (Giacomini, 2005).

For the decisions that set indication criteria; Mladovsky and Sorenson describe how the necessity of coverage is lowered when doctors rely "on patients to ascribe infertility" (Mladovsky & Sorenson, 2010). Argumentations with concrete, micro-level indication criteria include, for example, "the number of children living in the home/from previous relationships" (Brown, 2000; McWhirter & McQueen, 2000; Plomer et al., 1999; Shaw et al., 2002).

4. Viagra

Viagra is almost never covered (Klein & Sturm, 2002; Stolk et al., 2002), and moreover, about half of the argumentations oppose coverage. The other half contains both positive decisions and those that set indication criteria.

Viagra is, in actual coverage decisions, considered not medically necessary (Klein & Sturm, 2002; Stolk et al., 2002), patients not in need (Klein & Sturm, 2002), and erectile dysfunction not serious enough to prevent societal functioning (Hoedemaekers & Oortwijn, 2003) or to be a normal part of ageing (Stolk et al., 2002). Moreover, the apparent subjectivity of this determination of need by the patient (rather than the doctor) may have further loosened the concept of 'medical necessity' for Viagra. This may well effectuate over-usage, as policy makers feel patients cannot be fully relied upon to make this kind of decisions (Klein & Sturm, 2002). Moreover, because Viagra is available "over-the-web," the doctor is thus conclusively unneeded for diagnosis and prescription, which decreases the necessity of coverage, further strengthened by fear of moral hazard (Klein & Sturm, 2002).

Others, primarily authors, do support coverage of Viagra, based on the severity of psychosocial problems (Hornbrook & Holup, 2011; Klein & Sturm, 2002; Manson, 2005), impeded societal functioning as sexual function is considered essential to quality of life (Hornbrook & Holup, 2011) and positive coverage of similar treatments (Hornbrook & Holup, 2011).

Indication criteria are set in several countries, generally holding that Viagra as a treatment for erectile dysfunction is covered in cases of a spinal cord injury only, an argumentation related to the definition of illness and morbidity (Klein & Sturm, 2002; Manson, 2005; Stolk et al., 2002). This use of the definition of illness argumentation is also questioned once (Manson, 2005).

5. Cosmetic surgery

Both the public and policy makers often regard cosmetic surgical procedures like tattoo removal or 'boob jobs' a prime example of a form of care that should not be funded (Polikowski & Santos-Eggimann, 2002; Russell et al., 2014); in several countries it is indeed excluded from coverage (Schreyögg et al., 2005). The decisions examined all set limits on eligibility through indication criteria or, in cases like circumcision, exclude the form of care altogether.

The main argumentation is offered in support of a decision setting indication criteria, that is, cosmetic surgery is generally reimbursed for certain indications only. All decisions, which included primarily actual and a few hypothetical decisions, use certain physical characteristics, evidence of physical or psychosocial morbidity, reduced social capacity, or a combination thereof, to argue for necessity of coverage *in certain cases*. Decision makers thus set physical or psychological patient eligibility criteria on the morbidity experienced that need to be applied by a physician (Adler, 2011; Benditte-Klepetko et al., 2007; Breuning et al., 2010; Cook et al., 2003; Goodson et al., 2011; Henderson, 2009; Horner, 2002; Kerrigan et al., 2002; Krieger & Lesavoy, 2001; McClean & Hanke, 1997; Mukherjee et al., 2014; Nguyen et al., 2008; Russell et al., 2014; Stevens et al., 2015; A. M. Wong, 1995; Wraight et al., 2007). Individual assessment providing evidence of morbidity here leads to a higher justifiability of surgery. In this way, surgery is performed only on those patients that fall on the 'right' side of the line between cosmetic and non-cosmetic care (Goodson et al., 2011; Krieger & Lesavoy, 2001) or outside a "range of normality" (Breuning et al., 2010; Cook et al., 2003; Goodson et al., 2011; Henderson, 2009; Kerrigan et al., 2002; McClean & Hanke, 1997; Mukherjee et al., 2014; Nguyen et al., 2008; Nicoletti et al., 2009; Russell et al., 2014; Schnur et al., 1991; Stevens et al., 2015) so that care is only provided if medically necessary (Benditte-Klepetko et al., 2007; True, 2012). Some argue that cosmetic surgery should be covered only in case of societal functioning problems (Breuning et al., 2010; Henderson, 2009; Mukherjee et al., 2014), which McClean and Hanke argue to be affected only in some cases (McClean & Hanke, 1997). Directly related is the illness definition (Adler, 2011; Stevens et al., 2015): what counts as illness needs treatment and what falls inside the 'range of normality' does not need treatment.

This is also reflected in the argumentation supporting a negative coverage decision, stating that because "newborns do not have a medical condition," circumcision should not be covered (Adler, 2011), see also (Darby, 2016; Krieger & Lesavoy, 2001). Other argumentations against coverage

include, for actual decisions, the availability of alternative treatments (True, 2012) and the lack of medical necessity (McClean & Hanke, 1997), the latter also in hypothetical decisions (Adler, 2011; Darby, 2016; Jacobs, 1980).

Argumentations for a hypothetical positive decision were offered primarily for treatment of port-wine stains, with medical complications and psychosocial morbidity explicated (McClean & Hanke, 1997), also (Henderson, 2009). The (potential of a) 'postcode lottery' was brought up a number of times, to apparently argue in favour of coverage (Goodson et al., 2011; Henderson, 2009; Mukherjee et al., 2014; Stevens et al., 2015; Wraight et al., 2007).

6. Obesity treatments

Obesity treatments are often excluded from coverage (Polikowski & Santos-Eggimann, 2002). However, various argumentations in support of coverage are visible in the literature and a few in combination with indication criteria. In some cases, argumentations against coverage are given, but there is no evidence of public outrage over these negative decisions (Greer et al., 2012; L. F. Martin et al., 1998; Persson et al., 2010; Simpson & Cooper, 2009).

Articles mention the psychosocial burden of obesity (L. F. Martin et al., 1998; Persson et al., 2010) as supporting hypothetical coverage decisions, as well as the fact patients should be enabled to "conform to societal standards," which is linked to the range of normality mentioned for cosmetic surgery. For actual decisions, potential co-morbidities (Persson et al., 2010), as well as the lack of alternatives and simply "need and solidarity principles" (Persson et al., 2010) are brought up. Further hypothetical argumentations favouring coverage include societal responsibility for these patients (Greer et al., 2012), "fair treatment" (L. F. Martin et al., 1998), consideration of "human dignity" (Persson et al., 2010), and "social justice" (Greer et al., 2012).

Clear argumentations in support of decisions that set indication criteria are visible in decisions by American insurance companies, with positive decisions based on an expert opinion (L. F. Martin et al., 1998), the medical necessity as judged by a doctor (Simpson & Cooper, 2009) or a health risk assessment (Greer et al., 2012).

Conversely, the argumentations opposing coverage used by insurance companies in the USA hold that the individual patient is "to blame," and that obesity treatment should therefore not be included (Greer et al., 2012; L. F. Martin et al., 1998) but also that obesity is not an illness (L. F. Martin et al., 1998) or not severe enough (Simpson & Cooper, 2009).

7. Smoking cessation therapies

Smoking cessation therapy is covered in countries such as Australia, Canada, England, and the USA, though not in Argentina (Specogna, 2010), and it has been included and excluded several times in the Netherlands (Kroneman & de Jong, 2015). The subset contains negative and positive coverage decisions and focuses on the question of societal versus individual responsibility, only observed once before (in the obesity treatment subset).

For negative coverage decisions, two argumentations were elucidated. First, smoking or the cessation of it is considered in actual decisions to be a "personal" issue (Au-Yeung et al., 2010; Heath et al., 2002). Second, "tobacco consumption [is] not understood as an addiction or illness," which reduces the necessity of coverage.

For positive coverage decisions, the four argumentations include first, for actual decisions, the obligation to attempt "safe-guarding the wellbeing of fund participants" (Au-Yeung et al., 2010). Second, for hypothetical decisions, it is argued that distributive rather than retributive justice ought to prevail (Heath et al., 2002). Third, addiction is an illness (Heath et al., 2002; Woolf et al., 2006), also in use in actual decisions (Kroneman & de Jong, 2015). The fourth, though only used in hypothetical decisions, simply underlines the "need" for or "necessity" of smoking cessation services (Au-Yeung et al., 2010; Heath et al., 2002; Khalid, 1993; Specogna, 2010). The Dutch policy for coverage of smoking cessation therapy deserves special attention (Kroneman & de Jong, 2015). Smoking cessation therapy was excluded supported by the argumentation that the costs "could easily be paid from the savings that quitting smoking produced for individuals."

Analysis

In order to identify context-mechanism-outcome patterns (Pawson et al., 2005), we classified the twenty argumentation types per decision outcome: positive, negative, and conditional coverage (Table 3). First, seven argumentation types are generally used in the justification of positive coverage decisions (both hypothetical and actual), namely: Equity/Fairness/Justice, Societal Functioning, Individual Cost, the coverage of Similar Treatments, Human Rights, Dignity, and Societal Impact. Second, five argumentation types are solely in use in the justification of negative coverage decisions (both hypothetical and actual): Medical Necessity, Individual Responsibility, Moral Hazard, Small Number of Patients, and Patient-Diagnosis. Finally, eight argumentation types are employed for all types of coverage decision outcomes. This set comprises Morbidity/Severity, Range of Normality, Definition of Illness, Vulnerability/Compassion, Need, (No) Alternatives, Societal Responsibility, and the Rule of Rescue. Conditional coverage is, moreover, generally based on Morbidity/Severity, Definition of Illness, Range of Normality, Societal Functioning, and Medical Necessity – doctors are in these cases may make the final individual coverage decision. From Table 3, we thus observe that nearly two-thirds of the argumentation types is in use for either positive or negative decisions, from which we conclude the **first meta-level theory**: different argumentation types are generally used for different decision outcomes.

Table 4 gives a précis of those decisions where policy makers made a coverage decision responding to a patient (representative), the media, or the public, and it was explicitly analysed in the article as such. These actors were thus reported to be involved in setting the decision on the agenda, often reacting to a negative decision. We quickly detected that all coverage decisions in this table have a positive outcome. We first identified two decisions (Aggarwal et al., 2014; Burls et al., 2005) that use an argumentation type, in both cases Equity/Fairness/Justice, which is generally used in positive decisions (as visible in Table 3), meaning that the positive outcome was to be expected from the

Mechanism	Outcome					
Argumentation type	POSITIVE COVERAGE		NEGATIVE COVERAGE		CONDITIONAL COVERAGE	
	Actual	Hypothetical	Actual	Hypothetical	Actual	Hypothetical
Equity/Fairness/ Justice		Cancer; Cosmetic; Infertility; Obesity; Orphan; Smoking	Orphan			
Societal Functioning	Infertility	Cancer; Cosmetic; Infertility; Obesity; Orphan; Viagra	Viagra		Cosmetic	
Individual Cost		Infertility; Orphan	Smoking			
Similar Treatments		Infertility; Viagra	Orphan			
Human Right		Infertility				
Dignity		Cancer; Obesity				
Societal Impact		Cancer; Infertility				
Medical Necessity		Smoking	Cosmetic; Infertility; Viagra	Cosmetic; Infertility	Obesity	Cosmetic
Individual Responsibility			Infertility; Smoking	Obesity		
Moral Hazard				Infertility; Viagra		
Small Number of Patients	Cancer			Cancer; Orphan		
Patient-Diagnosis				Viagra		
Morbidity/Severity	Cancer; Infertility; Viagra	Cancer; Cosmetic; Infertility; Obesity; Orphan; Viagra	Cosmetic; Obesity; Viagra	Cosmetic; Infertility; Orphan	Cosmetic; Infertility; Orphan; Viagra	
Range of Normality		Cosmetic; Infertility	Viagra	Cosmetic	Cosmetic; Infertility	Cosmetic
Definition of Illness	Obesity; Smoking	Infertility; Smoking; Viagra	Infertility; Obesity; Smoking	Infertility	Cosmetic; Infertility; Obesity; Viagra	Cosmetic
Vulnerability/ Compassion	Cancer; Orphan	Cancer; Cosmetic; Infertility; Orphan	Cosmetic	Infertility		
Need	Cancer; Infertility; Obesity	Cancer; Cosmetic; Infertility; Smoking	Viagra	Infertility; Orphan		
[No] Alternative	Cancer; Obesity	Cancer; Infertility; Orphan	Cosmetic; Infertility	Infertility		
Societal Responsibility	Smoking	Obesity	Infertility	Infertility		Smoking
Rule of Rescue		Cancer; Orphan		Cancer; Orphan	Orphan	

Table 3 - Articles containing argumentations used to justify coverage decisions per treatments or therapies. 'Cancer' stands for cancer therapies, 'orphan' for orphan drugs, 'infertility' for infertility treatments, 'cosmetic' for cosmetic surgery, 'obesity' for obesity treatment and 'smoking' for smoking cessation therapy. NB: When a particular argumentation-outcome combination occurred only in one article, the cell is coloured light grey.

Context		Mechanism	Outcome
Country	Actor	Argumentation type(s)	Decision
UK	Politician	Equity/Fairness/Justice	Positive
Canada	Patient representative	Equity/Fairness/Justice	Positive
Canada	Patient representative	Rule of Rescue	Positive
Australia	Journalist	Rule of Rescue Small Number of Patients No Alternative	Positive
USA	Patient	Morbidity/Severity Medical Necessity Societal Functioning	Positive
UK	Patient	(not explicated, possibly Morbidity/Severity or Need)	Positive
Netherlands	Lobbyist	(not explicated, possibly Societal Responsibility or Societal Impact)	Positive

Table 4 - Overview of decisions where it was recorded that the public, a patient or a patient representative was involved in setting the coverage decision on the policy agenda, with the argumentation type and final decision

argumentation type used. However, we also found five cases that employ argumentation types also sometimes used to justify negative decisions, but that had, in fact, also a positive outcome (Burls et al., 2005; Clarke et al., 2001; Kroneman & de Jong, 2015; MacKenzie et al., 2008; McClean & Hanke, 1997). From this, we conclude that who set the decision on the agenda may have played a role in these cases. The **second meta-level theory** is therefore: when the public or a patient (representative) is recorded to have been involved in setting the decision agenda, regardless of the argumentation type used, the decision outcome is positive in this dataset. Indubitably, we would need a more extensive analysis of grey literature and media coverage to 'prove' this conclusively or to draw conclusions for unexamined coverage decision processes. It seems conceivable, however, that the working of the argumentation types depends upon the context. We conclude that an important context is who set the agenda; when a patient (representative), lobbyist, journalist, or politician was recorded to be involved, the resulting decision was always positive.

For the latter two meta-level theories, we utilise the distinction between 'actual' and 'hypothetical' coverage decisions. Actual decisions are those made by policy makers and insurance companies, that is, they concern coverage policies and therefore directly affect patients' access to medicine or therapy. Hypothetical decisions are those made by all other parties, which includes patients, authors (generalising, primarily ethicists and economists), the general public, and the media. We grouped the twenty necessity argumentation types into two broad categories, which are based on the type of decision (actual vs. hypothetical) they are used in. The categorisation is then made as follows: the first category is argumentations used in *both* actual and hypothetical decisions, the second category contains argumentations *only* used in hypothetical decisions.

The first category (as visible in the upper half of Table 5) contains ten argumentation types: Definition of Illness, Individual Responsibility, Medical Necessity, Morbidity/Severity, Need, (No) Alternative, Range of Normality, Societal Functioning, Societal Responsibility, and Vulnerability/ Compassion. The lower half of Table 5 holds the second category of necessity argumentations,

Mechanism		Outcome	
		Decision type	
Argumentation type		Actual decision [Policy makers and insurance companies]	Hypothetical decision [Patients, authors, the public, and the media]
Used in both actual and hypothetical decisions	Definition of Illness	Cosmetic; Infertility; Obesity; Smoking; Viagra	Cosmetic; Infertility; Smoking; Viagra
	Individual Responsibility	Infertility; Smoking	Obesity
	Medical Necessity	Cosmetic; Infertility; Obesity; Viagra	Cosmetic; Infertility; Smoking
	Morbidity/Severity	Cancer; Cosmetic; Infertility; Obesity; Orphan; Viagra	Cancer; Cosmetic; Infertility; Obesity; Orphan; Viagra
	Need	Cancer; Infertility; Obesity; Viagra	Cancer; Cosmetic; Infertility; Orphan; Smoking
	[No] Alternative	Cancer; Cosmetic; Infertility; Obesity	Cancer; Infertility; Orphan
	Range of Normality	Cosmetic; Infertility; Viagra	Cosmetic; Infertility
	Societal Functioning	Cosmetic; Infertility; Viagra	Cancer; Cosmetic; Infertility; Obesity; Orphan; Viagra
	Societal Responsibility	Smoking; Infertility	Infertility; Obesity; Smoking
	Vulnerability/Compassion	Cancer; Cosmetic; Orphan	Cancer; Cosmetic; Infertility; Orphan
Used in hypothetical decisions only	Dignity		Cancer; Obesity
	Equity/Fairness/Justice	Orphan	Cancer; Cosmetic; Infertility; Obesity; Orphan; Smoking
	Human Rights		Infertility
	Individual Cost	Smoking	Infertility; Orphan
	Moral Hazard		Infertility; Viagra
	Patient-Diagnosis		Viagra
	Rule of Rescue	Orphan	Cancer; Orphan
	Similar Treatments	Orphan	Infertility; Viagra
	Small Number of Patients	Cancer	Cancer; Orphan
	Societal Impact		Cancer; Infertility

Table 5 - Twenty argumentation sets organised per category, and the articles containing argumentations per treatment set per decision type of outcome (actual vs. hypothetical). 'Cancer' stands for cancer therapies, 'orphan' for orphan drugs, 'infertility' for infertility treatments, 'cosmetic' for cosmetic surgery, 'obesity' for obesity treatment and 'smoking' for smoking cessation therapy. NB: When a particular argumentation-outcome combination occurred only in one article, the cell is coloured light grey.

which are only used in hypothetical decisions. It consists of the following ten argumentation types: Dignity, Equity/Fairness/Justice, Human Rights, Individual Cost, Moral Hazard, Patient-Diagnosis, Rule of Rescue, Similar Treatments, Small Number of Patients, and Societal Impact. Furthermore, we observe that the first category contains the three most strongly acknowledged argumentation types, namely Morbidity/Severity (acknowledged thirteen times), (No) Alternative (eight times), and Need (seven times). In contrast, the second category holds the two most strongly questioned argumentations (that is, more than once), namely Small Number of Patients (six times) and Rule of Rescue (four times), though these were also as often acknowledged as valid (seven times and five times, respectively). Based on this analysis, we formulate the **third meta-level theory**: half of

Mechanism		Context				
Argumentation type		Australia	Canada	USA	UK	France, Germany, Sweden, the Netherlands
Used in both actual and hypothetical decisions	Definition of Illness	-	4.9	19.2	5.6	13.5
	Individual Responsibility	-	-	5.8	1.1	-
	Medical Necessity	-	4.9	19.2	-	2.7
	Morbidity/Severity	25.0	19.5	19.2	28.1	16.2
	Need	-	7.3	5.8	7.9	18.9
	(No) Alternatives	25.0	12.2	5.8	7.9	16.2
	Range of Normality	-	-	7.7	16.9	5.4
	Societal Functioning	-	2.4	3.8	5.6	2.7
	Societal Responsibility	-	-	1.9	-	2.7
	Vulnerability/Compassion	16.7	12.2	3.8	4.5	10.8
Used in hypothetical decisions only	Dignity	-	2.4	-	-	2.7
	Equity/Fairness/Justice	-	7.3	5.8	13.5	-
	Individual Cost	-	-	-	-	2.7
	Moral Hazard	-	-	1.9	-	-
	Rule of Rescue	25.0	12.2	-	1.1	-
	Similar Treatments	-	2.4	-	2.2	-
	Small Number of Patients	8.3	9.8	-	5.6	5.4
	Societal Impact	-	2.4	-	-	-
	Total number of argumentations	12	41	52	89	37

Table 6 - Percentages of use of argumentations per country (for which a minimum of five argumentations was found). The argumentations are shown in alphabetical order per category. Human Rights and Patient-Diagnosis are absent as they were not used in country-specific argumentations.

the argumentation types are used by policy makers and insurance companies as well as patients, authors, the public, and the media, whereas the other half are only used by the latter group. The argumentation types that are used by policy makers and insurance companies are, moreover, more frequently acknowledged as valid, whereas some that are only used by patients, authors, the public, and the media are strongly questioned.

In an effort to understand whether and how the geographical location (the second context that we observed the impact of) affected the use of argumentation types, we plotted the usage frequency per country in Table 6. We identified the following parallels between the countries. From the first necessity category (the upper half of Table 6, categorised by use as described above), most countries employ most argumentation types at some point, and all countries employ the argumentation types Morbidity/Severity, No Alternatives, and Vulnerability/Compassion. In contrast, Medical Necessity is constricted to the USA, and Need is most frequently visible for 'mainland Europe' (France, Germany, Sweden, and the Netherlands). For the second necessity category (the lower half of Table 6), the pattern is much less congruent. Canada and Australia have the highest percentage of use

of argumentations grouped into this category. Moreover, only these countries utilise the more controversial argumentation types, namely Rule of Rescue and Small Number of Patients, whereas this is not the case for the other countries. This is to be expected, as the Rule of Rescue is part of the official criteria of the PBAC (cf. Cookson et al., 2008). In addition, the UK has a high percentage of the Equity/Fairness/Justice argumentation type, which is primarily due to the concern and outrage about 'postcode rationing' (Breuning et al., 2010; Goodson et al., 2011; Henderson, 2009; Mukherjee et al., 2014; Stevens et al., 2015; Wraight et al., 2007).

Accordingly, we formulated the **fourth and final meta-level theory**, which holds that the argumentations that fall under the first necessity category appear to be applied in most of the (predominantly Western) countries the dataset reports on, though the countries do appear to have their own argumentation type preferences. For the second category the patterns are much more varied: many argumentation types have a specific pattern (that is, used in two or three countries), which we conclude as depending on the local context.

DISCUSSION

In this chapter we unpack the question of necessity of coverage by reviewing the argumentations used and proposed to be used in coverage decision making. In seven treatment sets (cancer therapies, orphan drugs, infertility treatments, Viagra, cosmetic surgery, obesity treatment, and smoking cessation therapy) from 98 reviewed articles, we identified twenty different argumentation types employed to argue for or against the necessity of coverage of a certain therapy or treatment.

The argumentation types may be typified by primary decision outcome. The following eight argumentation types are used primarily in favour of coverage: equity, fairness and justice, a patient's societal functioning being impeded, the cost being too high to bear for an individual, the coverage status of similar treatments, considerations concerning human rights and dignity of the patient, and wider societal impact. Reversely, limited medical necessity, the individual being held responsible, potential moral hazard or over-usage, a small patient population, and a doctor needing to rely on the patient to set the diagnosis are argumentations solely employed to decrease the necessity of coverage. Finally, the morbidity experienced or the severity of the disease, whether the ailment falls within a defined 'range of normality' or is considered an illness, compassion for vulnerable patients, the need of the patient, the (un)availability of alternatives, society having to take responsibility for coverage, and the rule of rescue are used both in favour of and opposing coverage. The most important context affecting the argumentations in use in coverage decisions is the way the decision has come onto the agenda: when this is recorded in the article to be by a patient, a patient organisation, or the public, this always facilitates a positive decision, even when the argumentation type employed is also in use in negative coverage decisions. Others have stressed much the same point (Booth et al., 2007); in a cross-country comparison of hepatitis C coverage, Kieslich et al. describe experiences that are "as much a tale of challenges that arise when making difficult prioritization

decisions as they are a tale of agenda-setting" (Kieslich et al., 2016). Overall, it is important to note that no wide consensus is needed to let a consideration function as an argumentation in coverage decisions. This makes it doubly interesting to investigate what exactly happens during the processes of coverage decision making.

Necessity is also broadly categorised by two different types of usage. Under the first necessity category, indicating use in both actual and hypothetical decisions, fall the definition of illness and range of normality, the individual's own responsibility, the medical necessity, the morbidity (or the severity of the disease) and need experienced, the presence or absence of alternatives, the societal functioning of the patient, the societal responsibility for coverage, and compassion towards vulnerable patients. The second necessity category is not used in actual coverage decisions, that is, by policy makers and insurance companies but rather only by patients, authors, the public, and the media. For this second necessity category, we observed dignity, equity, fairness and justice considerations, human rights, the individual cost of the treatment, moral hazard, whether the patient sets the diagnosis him- or herself, the rule of rescue, coverage of similar treatments, the size of the patient population, and the societal impact of coverage. Interestingly, the use of the argumentation types that fall under this category is more localised, that is, country-specific, and more frequently debated. The argumentation types belonging to the first necessity category, however, are much more universally applied.

Comparison with existing literature

Our search string yielded seven studies containing literature reviews that were included in our review (Aggarwal et al., 2014; Drummond et al., 2014; Mladovsky & Sorenson, 2010; Nguyen et al., 2008; Paulden et al., 2015; Rocchi et al., 2008; Zelei et al., 2016), as well as a number of reviews that specified general qualitative criteria for coverage decisions. Cromwell, Peacock, and Mitton have conducted a literature review of 'real-world' decision criteria as found in 33 articles in peer-reviewed and grey literature (Cromwell et al., 2015). They searched for articles with an explicit acknowledgement of a "finite resource pool." Therefore, their review focuses on resource allocation, i.e., choosing one thing over another, in different settings such as hospitals and health authorities. They found that disease impact (burden) was an influential criterion in resource allocation decisions but did not observe other necessity-based argumentations.

Several other studies, e.g., Cerri et al., examined technology appraisals as conducted by macro-level institutions and commonly find it to be a "complex process involving numerous clinical, disease, and affordability considerations" (Cerri et al., 2014). Shah examined popular preferences in using severity of illness in economic evaluation as part of priority setting, showing that "people are, on the whole, willing to sacrifice aggregate health in order to give priority to the severely ill," which may further legitimise the use of individual burden of disease as a coverage criterion (Shah et al., 2014). Fischer conducted a broad review of decisions where quantitative methods were employed and identified several clinical criteria that strengthen our conclusions: availability of treatment alternative; condition is life threatening; condition caused by patients own behaviour; and end

of life considerations (Fischer, 2012). Hasman, McIntosh, and Hope, finally, interviewed decision makers on what they considered relevant in a coverage decision on a hypothetical drug (Hasman et al., 2008). The interviews highlighted agreement over cost effectiveness, clinical effectiveness, equality, and gross cost as important reasons for coverage. Interestingly, a lack of agreement was elicited over using 'absence of alternative treatments' as a criterion. Our study, rather than aiming to outline general criteria in use in coverage decisions, specifically gives an analysis of the tonal variety in the use of necessity in actual and hypothetical coverage decisions, and the effect contexts may have on the final coverage decision.

Strengths, limitations, and future research directions

To study, like we have, the way contexts influence coverage decision outcomes through the use of necessity as a criterion will aid our understanding of both coverage decision making as well as necessity as a concept as used in practice. In addition, the wide variety of argumentations described will aid further ethical reflection on the content and processes of actual and hypothetical coverage decisions.

Our search strategy in particular provided a challenge, as when one searches for "need" in the title/abstract, the articles that will be found are likely to discuss coverage decisions that use "need" as a criterion. As such, certain necessity-based argumentations were in that sense already to be expected from the wording of the search strategy. We have engaged two expert librarians to aid us in order to widen and strengthen our search terms and have thus acquired a relatively 'broad but clean' dataset.

The use of the realist review method has sharpened our review by forcing us to both broaden our choice in articles (opinion pieces, ethical and economic analyses, and reviews were analysed on top of document analyses and interviews/surveys) and take the defined contexts, rather than the practical set-up of the research, as primary quality attribute for inclusion and scoring. Although not always easy to accomplish, this study shows that the realist review method may successfully be applied to a wider variety of interventions than classical policy interventions such as patient portals (Otte-Trojel et al., 2014) or internet-based medical education (Wong et al., 2010). Necessity as a criterion is a policy 'product' that interacts with a variety of contexts to produce certain outcomes, and to frame our question thus has elicited exciting data on the varied use of necessity as a criterion in coverage decisions from peer-reviewed literature. It must, however, also be stressed that we have not effectuated a different 'weighing' of actual versus hypothetical decisions and that ethical discussions and public attitude surveys are of a wholly different substance than coverage decisions made by policy makers as they directly affect patients' lives. The inclusion of hypothetical decisions has not been self-evident, but it has been crucial to giving an overview of as wide a selection of criteria-in-action as possible.

It is, however, debatable to what extent we have grasped the full context of these criteria-in-action due to the way 'context' has been defined in this study. For example, on the micro level in Sweden, despite the national-level guidance not to treat smokers differently, "physicians are more inclined to

treat a non-smoking patient" (Björk et al., 2015). The element of weighing personal responsibility in coverage decisions is visible here but rarely expressed as such in health policy (Tinghög et al., 2010). This shows that criteria may be exhibited in other argumentations at the bedside (e.g., (Breuning et al., 2010; Rooshenas et al., 2015) or in court (e.g., (Giacomini, 2005; McIver & Ham, 2000). We have also omitted the characteristics of the exact form of care under consideration as being potentially influential, insofar as they were not described as part of the argumentations (cf. Lowi in (Kieslich et al., 2016). Additionally, one could argue that the way contexts have been defined in this study obscures the role of the decision maker or the structure of the decision-making process, as these also have not been considered as 'contexts'. Furthermore, the 'agenda setting' context is a short consideration of decisions that may have been made for political, rather than explicitly patient-focused, reasons (Gordon, 2006; Kroneman & de Jong, 2015; Specogna, 2010). Lastly, the data contained few articles on non-Western countries, which, upon becoming available, may well nuance the homogeneity in use of necessity considerations between countries. This, as well as additional research on the contexts mentioned earlier, may prove fruitful lines of further inquiry.

Finally, it would appear that using the realist review method predisposes one to think in terms of causality: a certain mechanism in a certain context leads to a certain outcome; argumentation type M, when used in country C leads to (hypothetical or actual) inclusion of treatment T. This, in itself, is problematic: the fact that a combination occurred 'out there' does not necessarily mean it *needed* to happen that way; it underplays the agency of the decision makers in this important sense. It also assumes that the outcome 'comes last'; it leaves no space for (additional) justifications that may have been brought in after the decision was made. Further investigation, for example through an ethnography, of the coverage decision making process and the role and use of necessity as a criterion in coverage decisions would be a useful exercise to shed more light hereon.

CONCLUSION

Necessity as a criterion in coverage decisions has lacked operationalisation: this chapter provides a handle on the wide variety of argumentation types that fall under this umbrella term by reviewing over 400 argumentations described in 98 peer-reviewed articles. These argumentations are grouped into twenty different necessity-based argumentation types that are used in coverage decisions, both 'actual' decisions (made by policy makers and insurance companies) as well as 'hypothetical' decisions (made by actors such as patients, authors, the public, and the media).

Eight of these twenty necessity-based argumentations are used to support both positive and negative coverage decisions; twelve are, however, only employed for either positive or negative decisions. When patients or the public were recorded to have aided in setting the decision on the agenda this always resulted in a positive decision, even when an argumentation type was used that could have resulted in a negative decision.

The argumentation types heightening or lowering necessity of coverage of a certain form of care may also be distinguished by their users. Half of the argumentation types is used in both actual and hypothetical coverage decisions, that is, by policy makers and insurance companies as well as patients, authors, the public, and the media. The second half is, in contrast, only used by the latter group. Argumentation types used in actual as well as hypothetical decisions are more strongly acknowledged and used in a greater number of countries. Conversely, argumentation types used in hypothetical decisions only are much more strongly questioned, and their use is much more country specific.

REFERENCES

Adler, P. W. (2011). Is It Lawful to Use Medicaid to Pay for Circumcision? J Law Med, 19(2), 335-353.

Aggarwal, A., Ginsburg, O., & Fojo, T. (2014). Cancer Economics, Policy and Politics: What Informs the Debate? Perspectives from the Eu, Canada and Us. Journal of Cancer Policy, 2(1), 1-11. doi: https://doi.org/10.1016/j.jcpo.2014.02.002

Akhmetov, I., & Bubnov, R. V. (2015). Assessing Value of Innovative Molecular Diagnostic Tests in the Concept of Predictive, Preventive, and Personalized Medicine. EPMA J, 6(1). doi: 10.1186/s13167-015-0041-3

Au-Yeung, C. M., Weisman, S. R., Hennrikus, D. J., Forster, J. L., Skoog, R., Luneburg, W., & Hesse, B. (2010). An Evaluation of Health Benefit Modification in Taft-Hartley Health and Welfare Funds: Implications for Encouraging Tobacco-Cessation Coverage. Am J Prev Med, 39(6 SUPPL. 1), S30-S36. doi: 10.1016/j.amepre.2010.09.010

Bae, G., Bae, E. Y., & Bae, S. (2015). Same Drugs, Valued Differently? Comparing Comparators and Methods Used in Reimbursement Recommendations in Australia, Canada, and Korea. Health Policy, 119(5), 577-587.

Bañón Hernández Antonio, A. (2015). The Debate on Rare Diseases: A Look at Media Response. Mètode Science Studies Journal: Annual Review, 0(6).

Benditte-Klepetko, H., Leisser, V., Paternostro-Sluga, T., Rakos, M., Trattnig, S., Helbich, T., . . . Deutinger, M. (2007). Hypertrophy of the Breast: A Problem of Beauty or Health? J Women's Health, 16(7), 1062-1069. doi: 10.1089/jwh.2006.0183

Bernardi, A., & Pegoraro, R. (2003). Italian Drug Policy: Ethical Aims of Essential Assistance Levels. Health Care Anal, 11(4), 279-286.

Bernfort, L. (2003). Decisions on Inclusion in the Swedish Basic Health Care Package - Roles of Cost-Effectiveness and Need. Health Care Anal, 11(4), 301-308. doi: 10.1023/B:HCAN.0000010059.61453.8e

Björk, J., Lynøe, N., & Juth, N. (2015). Are Smokers Less Deserving of Expensive Treatment? A Randomised Controlled Trial That Goes Beyond Official Values. BMC medical ethics, 16(1), 1.

Blank, R. H. (1997). Assisted Reproduction and Reproductive Rights: The Case of in Vitro Fertilization. Politics Life Sciences, 16(2), 279-288.

Booth, C. M., Dranitsaris, G., Gainford, M. C., Berry, S., Fralick, M., Fralick, J., . . . Clemons, M. (2007). External Influences and Priority-Setting for Anti-Cancer Agents: A Case Study of Media Coverage in Adjuvant Trastuzumab for Breast Cancer. BMC Cancer, 7(1), 110. doi: 10.1186/1471-2407-7-110

Bramer, W., Holland, L., Mollema, J., Hannon, T., & Bekhuis, T. (2014). Removing Duplicates in Retrieval Sets from Electronic Databases: Comparing the Efficiency and Accuracy of the Bramer-Method with Other Methods and Software Packages.

Breuning, E. E., Oikonomou, D., Singh, P., Rai, J. K., & Mendonca, D. A. (2010). Cosmetic Surgery in the Nhs: Applying Local and National Guidelines. Journal of Plastic, Reconstructive & Aesthetic Surgery, 63(9), 1437-1442.

Brock, D. W. (2010). Ethical and Value Issues in Insurance Coverage for Cancer Treatment. Oncologist, 15 Suppl 1, 36-42. doi: 10.1634/theoncologist.2010-S1-36

Brown, C. J. (2000). Rationing Fertility Services in the Nhs: The Patients' Viewpoint. Hum Fertil, 3(4), 268-270.

Burls, A., Austin, D., & Moore, D. (2005). Commissioning for Rare Diseases: View from the Frontline. Bmj, 331(7523), 1019-1021. doi: 10.1136/bmj.331.7523.1019

Cerri, K. H., Knapp, M., & Fernandez, J. L. (2014). Public Funding of Pharmaceuticals in the Netherlands: Investigating the Effect of Evidence, Process and Context on Cvz Decision-Making. Eur J Health Econ, 15(7), 681-695. doi: 10.1007/s10198-013-0514-z

Chabot, I., & Rocchi, A. (2010). How Do Cost-Effectiveness Analyses Inform Reimbursement Decisions for Oncology Medicines in Canada? The Example of Sunitinib for First-Line Treatment of Metastatic Renal Cell Carcinoma. Value

Health, 13(6), 837-845. doi: 10.1111/j.1524-4733.2010.00738.x

Chambers, G. M., Adamson, G. D., & Eijkemans, M. J. C. (2013). Acceptable Cost for the Patient and Society. Fertil Steril, 100(2), 319-327. doi: 10.1016/j.fertnstert.2013.06.017

Chambers, G. M., Ho, M. T., & Sullivan, E. A. (2006). Assisted Reproductive Technology Treatment Costs of a Live Birth: An Age-Stratified Cost-Outcome Study of Treatment in Australia. Med J Aust, 184(4), 155-158.

Clarke, J. T. R., Amato, D., & Deber, R. B. (2001). Managing Public Payment for High-Cost, High-Benefit Treatment: Enzyme Replacement Therapy for Gaucher's Disease in Ontario. CMAJ: Canadian Medical Association Journal, 165(5), 595-596.

Cohen, J. P., & Felix, A. (2014). Are Payers Treating Orphan Drugs Differently? Journal of Market Access & Health Policy, 2, 10.3402/jmahp.v3402.23513. doi: 10.3402/jmahp.v2.23513

Connolly, M. P., Ledger, W., & Postma, M. J. (2010). Economics of Assisted Reproduction: Access to Fertility Treatments and Valuing Live Births in Economic Terms. Hum. Fertil., 13(1), 13-18.

Cook, S. A., Rosser, R., Meah, S., James, M. I., & Salmon, P. (2003). Clinical Decision Guidelines for Nhs Cosmetic Surgery: Analysis of Current Limitations and Recommendations for Future Development. Br J Plast Surg, 56(5), 429-436.

Cookson, R., Mccabe, C., & Tsuchiya, A. (2008). Public Healthcare Resource Allocation and the Rule of Rescue. J Med Ethics, 34(7), 540-544.

Couwenbergh, B., Van Der Meer, F., Weghaus-Reus, S., Schelleman, H., & Zwaap, J. (2013). Pakketbeheer in De Praktijk Deel 3. Diemen: Zorginstituut Nederland, 1-106.

Cromwell, I., Peacock, S. J., & Mitton, C. (2015). 'Real-World'health Care Priority Setting Using Explicit Decision Criteria: A Systematic Review of the Literature. BMC Health Serv. Res., 15(1), 1.

Daniels, N., & Sabin, J. E. (1998). Last Chance Therapies and Managed Care. Pluralism, Fair Procedures, and Legitimacy. Hastings Cent Rep, 28(2), 27-41.

Darby, R. (2016). Targeting Patients Who Cannot Object? Re-Examining the Case for Non-Therapeutic Infant Circumcision. SAGE Open, 6(2). doi: 10.1177/2158244016649219

Denis, A., Mergaert, L., Fostier, C., Cleemput, I., Hulstaert, F., & Simoens, S. (2011). Critical Assessment of Belgian Reimbursement Dossiers of Orphan Drugs. PharmacoEconomics, 29(10), 883-893. doi: 10.2165/11585980-000000000-00000

Desser, A. S. (2013). Prioritizing Treatment of Rare Diseases: A Survey of Preferences of Norwegian Doctors. Soc Sci Med, 94. doi: 10.1016/j.socscimed.2013.06.019

Desser, A. S., Gyrd-Hansen, D., Olsen, J. A., Grepperud, S., & Kristiansen, I. S. (2010). Societal Views on Orphan Drugs: Cross Sectional Survey of Norwegians Aged 40 to 67. Bmj, 341, c4715.

Drummond, M., De Pouvourville, G., Jones, E., Haig, J., Saba, G., & Cawston, H. (2014). A Comparative Analysis of Two Contrasting European Approaches for Rewarding the Value Added by Drugs for Cancer: England Versus France. PharmacoEconomics, 32(5), 509-520.

Drummond, M., Evans, B., Lelorier, J., Karakiewicz, P., Martin, D., Tugwell, P., & Macleod, S. (2009). Evidence and Values: Requirements for Public Reimbursement of Drugs for Rare Diseases--a Case Study in Oncology. The Canadian journal of clinical pharmacology = Journal canadien de pharmacologie clinique, 16(2), e273-281; discussion e282-274.

Drummond, M., & Mason, A. R. (2007). European Perspective on the Costs and Cost-Effectiveness of Cancer Therapies. J Clin Oncol, 25(2), 191-195. doi: 10.1200/jco.2006.07.8956

Drummond, M., & Towse, A. (2014). Orphan Drugs Policies: A Suitable Case for Treatment. Eur J Health Econ, 15(4), 335-340. doi: 10.1007/s10198-014-0560-1

Fischer, K. E. (2012). A Systematic Review of Coverage Decision-Making on Health Technologies-Evidence from the Real World. Health Policy, 107(2-3), 218-230. doi: 10.1016/j.healthpol.2012.07.005

Fischer, K. E., & Rogowski, W. H. (2014). Funding Decisions for Newborn Screening: A Comparative

Review of 22 Decision Processes in Europe. Int. J. Environ. Res. Public Health, 11(5), 5403-5430.

Foy, R., So, J., Rous, E., & Scarffe, J. H. (1999). Perspectives of Commissioners and Cancer Specialists in Prioritising New Cancer Drugs: Impact of the Evidence Threshold. BMJ, 318(7181), 456-459.

Giacomini, M. (2005). One of These Things Is Not Like the Others: The Idea of Precedence in Health Technology Assessment and Coverage Decisions. Milbank Q, 83(2), 193-223.

Giacomini, M., Hurley, J., & Stoddart, G. (2000). The Many Meanings of Deinsuring a Health Service: The Case of in Vitro Fertilization in Ontario. SOC SCI MED, 50(10), 1485-1500. doi: 10.1016/s0277-9536(99)00394-9

Goodson, A., Khoda, B., & Nduka, C. (2011). Funding Criteria for Common Procedures: A Postcode Lottery in Nhs Plastic Surgery. Journal of Plastic, Reconstructive & Aesthetic Surgery, 64(3), 417-419.

Gordon, E. J. (2006). The Political Contexts of Evidence-Based Medicine: Policymaking for Daily Hemodialysis. SOC SCI MED, 62(11), 2707-2719. doi: 10.1016/j.socscimed.2005.11.024

Greenhalgh, T., Wong, G., Westhorp, G., & Pawson, R. (2011). Protocol-Realist and Meta-Narrative Evidence Synthesis: Evolving Standards (Rameses). BMC Medical Research Methodology, 11(1), 115.

Greer, A. G., Chapman, C., & Ryckeley, J. B. (2012). Moral Hazard or Morality for Health: An Ethical Debate on Insurance Coverage for the Obese. Bariatr Nurs Surg Patient Care, 7(1), 36-41. doi: 10.1089/bar.2012.9992

Gross, M. L. (2002). Ethics, Policy, and Rare Genetic Disorders: The Case of Gaucher Disease in Israel. Theoretical Medicine and Bioethics, 23(2), 151-170. doi: 10.1023/a:1020324332620

Ham, C. (1997). Priority Setting in Health Care: Learning from International Experience. Health Policy, 42(1), 49-66. doi: http://dx.doi.org/10.1016/S0168-8510(97)00054-7

Hasman, A., Mcintosh, E., & Hope, T. (2008). What Reasons Do Those with Practical Experience Use in Deciding on Priorities for Healthcare Resources? A Qualitative Study. J Med Ethics, 34(9), 658-663. doi: 10.1136/jme.2007.023366

Heath, J., Braun, M. A., & Brindle, M. (2002). Smokers' Rights to Coronary Artery Bypass Graft Surgery. JONAS Healthc Law Ethics Regul, 4(2), 32-35.

Henderson, J. (2009). The Plastic Surgery Postcode Lottery in England. Int J Surg, 7(6), 550-558. doi: 10.1016/j.ijsu.2009.09.004

Henschke, C. (2012). Provision and Financing of Assistive Technology Devices in Germany: A Bureaucratic Odyssey? The Case of Amyotrophic Lateral Sclerosis and Duchenne Muscular Dystrophy. Health Policy, 105(2-3), 176-184. doi: 10.1016/j.healthpol.2012.01.013

Hoedemaekers, R., & Oortwijn, W. (2003). Problematic Notions in Dutch Health Care Package Decisions. HEALTH CARE ANAL, 11(4), 287-294. doi: 10.1023/B:HCAN.0000010057.43321.b2

Hornbrook, M. C., & Holup, J. (2011). Insurance Coverage for Erectile Dysfunction Drugs. Clinical Pharmacology & Therapeutics, 89(1), 19-21. doi: 10.1038/clpt.2010.265

Horner, B. (2002). Breast Augmentation Should Be on the Nhs: A Discussion of the Ethics of Rationing. Ann R Coll Surg Engl, 84(2), 82-83.

Hughes, D. (2006). Rationing of Drugs for Rare Diseases, Editorial, PharmacoEconomics, p. 315. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=20503290&site=ehost-live

Hughes, D., Tunnage, B., & Yeo, S. (2005). Drugs for Exceptionally Rare Diseases: Do They Deserve Special Status for Funding? Qjm, 98(11), 829-836.

Hughes, E. G., & Giacomini, M. (2001). Funding in Vitro Fertilization Treatment for Persistent Subfertility: The Pain and the Politics. Fertil. Steril., 76(3), 431-442.

Iskrov, G. G., Raycheva, R. D., & Stefanov, R. S. (2013). Insight into Reimbursement Decision-Making Criteria in Bulgaria: Implications for Orphan Drugs. Folia Med (Plovdiv), 55(3-4), 80-86.

Jacobs, S. A. (1980). The Determination of Medical Necessity: Medicaid Funding for Sex-Reassignment Surgery. Case West Reserve Law Rev, 31(1), 179-209.

Juth, N. (2014). For the Sake of Justice: Should We Prioritize Rare Diseases? Health Care Analysis, 1-20. doi: 10.1007/s10728-014-0284-5

Kapiriri, L., Norheim, O. F., & Martin, D. K. (2009). Fairness and Accountability for Reasonableness. Do the Views of Priority Setting Decision Makers Differ across Health Systems and Levels of Decision Making? Soc. Sci. Med., 68(4), 766-773.

Kerrigan, C. L., Collins, E. D., Kim, H. M., Schnur, P. L., Wilkins, E., Cunningham, B., & Lowery, J. (2002). Reduction Mammaplasty: Defining Medical Necessity. Med. Decis. Mak., 22(3), 208-217.

Khalid, M. I. (1993). Access to Surgery for Smokers. Denying Treatment Is Indefensible. BMJ : British Medical Journal, 306(6889), 1408-1408.

Kieslich, K., Ahn, J., Badano, G., Chalkidou, K., Cubillos, L., Hauegen, R. C., . . . Wilson, J. (2016). Public Participation in Decision-Making on the Coverage of New Antivirals for Hepatitis C. Journal of Health Organization and Management, 30(5), 769-785. doi: doi:10.1108/JHOM-03-2016-0035

Kisser, A., Tüchler, H., Erdös, J., & Wild, C. (2016). Factors Influencing Coverage Decisions on Medical Devices: A Retrospective Analysis of 78 Medical Device Appraisals for the Austrian Hospital Benefit Catalogue 2008-2015. Health Policy, 120(8), 903-912. doi: http://dx.doi.org/10.1016/j.healthpol.2016.06.007

Klein, R., & Sturm, H. (2002). Viagra: A Success Story for Rationing? Health Aff (Millwood), 21(6), 177-187.

Krieger, L. M., & Lesavoy, M. A. (2001). Managed Care's Methods for Determining Coverage of Plastic Surgery Procedures: The Example of Reduction Mammaplasty. Plast Reconstr Surg, 107(5), 1234-1240.

Kroneman, M., & De Jong, J. D. (2015). The Basic Benefit Package: Composition and Exceptions to the Rules. A Case Study. HEALTH POLICY, 119(3), 245-251. doi: 10.1016/j.healthpol.2015.01.013

Largent, E. A., & Pearson, S. D. (2012). Which Orphans Will Find a Home? The Rule of Rescue in Resource Allocation for Rare Diseases. Hastings Cent Rep, 42(1), 27-34.

Lim, C. S., Lee, Y. G., Koh, Y., & Heo, D. S. (2014). International Comparison of the Factors Influencing Reimbursement of Targeted Anti-Cancer Drugs. BMC Health Serv. Res., 14.

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. doi: 10.1002/hec.2872

Lord, J., Shaw, L., Dobbs, F., & Acharya, U. (2001). A Time for Change and a Time for Equality - Infertility Services and the Nhs. Hum Fertil, 4(4), 256-260.

Mackenzie, R., Chapman, S., Salkeld, G., & Holding, S. (2008). Media Influence on Herceptin Subsidization in Australia: Application of the Rule of Rescue? J. R. Soc. Med., 101(6), 305-312. doi: 10.1258/jrsm.2008.070289

Manson, H. (2005). The Role of the /`Lifestyle/' Label and Negative Bias in the Allocation of Health Resources for Erectile Dysfunction Drugs: An Ethics-Based Appraisal. Int J Impot Res, 18(1), 98-103.

Martin, D., Giacomini, M., & Singer, P. (2002). Fairness, Accountability for Reasonableness, and the Views of Priority Setting Decision-Makers. Health Policy, 61(3), 279-290.

Martin, D., Pater, J. L., & Singer, P. A. (2001). Priority-Setting Decisions for New Cancer Drugs: A Qualitative Case Study. Lancet, 358(9294), 1676-1681.

Martin, L. F., White, S., & Lindstrom, W., Jr. (1998). Cost-Benefit Analysis for the Treatment of Severe Obesity. World J Surg, 22(9), 1008-1017.

Mccabe, C., Claxton, K., & Tsuchiya, A. (2005). Orphan Drugs and the Nhs: Should We Value Rarity? Br Med J, 331. doi: 10.1136/bmj.331.7523.1016

Mcclean, K., & Hanke, C. W. (1997). The Medical Necessity for Treatment of Port-Wine Stains. Dermatol. Surg., 23(8), 663-667.

Mciver, S., & Ham, C. (2000). Five Cases, Four Actors and a Moral: Lessons from Studies of Contested Treatment Decisions. Health Expect., 3(2), 114-124.

Mcmillan, J. (2001). Allocating Fertility Services by Medical Need. Human Fertility, 4(1), 11-13. doi: 10.1080/1464727012000199191

Mcwhirter, M., & Mcqueen, D. (2000). Rationing and Nhs Fertility Services: A Commissioning Perspective. Hum Fertil, 3(4), 271-274.

Mentzakis, E., Stefanowska, P., & Hurley, J. (2011). A Discrete Choice Experiment Investigating Preferences for Funding Drugs Used to Treat Orphan Diseases: An Exploratory Study. Health Econ Policy Law, 6(3), 405-433. doi: 10.1017/s1744133110000344

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.

Mladovsky, P., & Sorenson, C. (2010). Public Financing of Ivf: A Review of Policy Rationales. Health Care Anal, 18(2), 113-128.

Mukherjee, S., Kamat, S., Adegbola, S., & Agrawal, S. (2014). Funding for Postbariatric Body-Contouring (Bariplastic) Surgery in England: A Postcode Lottery. Plastic Surgery International, 2014, 6. doi: 10.1155/2014/153194

Nachtigall, R. D., Macdougall, K., Davis, A. C., & Beyene, Y. (2012). Expensive but Worth It: Older Parents' Attitudes and Opinions About the Costs and Insurance Coverage for in Vitro Fertilization. Fertil. Steril., 97(1), 82-87.

Neumann, P. J. (1997). Should Health Insurance Cover Ivf? Issues and Options. J HEALTH POLIT POLICY LAW, 22(5), 1215-1239.

Nguyen, J. T., Wheatley, M. J., Schnur, P. L., Nguyen, T. A., & Winn, S. R. (2008). Reduction Mammaplasty: A Review of Managed Care Medical Policy Coverage Criteria. Plast Reconstr Surg, 121(4), 1092-1100. doi: 10.1097/01.prs.0000304238.43482.9c

Nicod, E. (2016). Why Do Health Technology Assessment Coverage Recommendations for the Same Drugs Differ across Settings? Applying a Mixed Methods Framework to Systematically Compare Orphan Drug Decisions in Four European Countries. The European Journal of Health Economics, 1-16. doi: 10.1007/s10198-016-0823-0

Nicoletti, G., Scevola, S., & Faga, A. (2009). Is Breast Reduction a Functional or a Cosmetic Operation? Proposal of an Objective Discriminating Criterion. Journal of Plastic, Reconstructive & Aesthetic Surgery, 62(12), 1644-1646. doi: http://dx.doi.org/10.1016/j.bjps.2008.06.069

Otte-Trojel, T., De Bont, A., Rundall, T. G., & Van De Klundert, J. (2014). How Outcomes Are Achieved through Patient Portals: A Realist Review. Journal of the American Medical Informatics Association, 21(4), 751-757.

Paulden, M., Stafinski, T., Menon, D., & Mccabe, C. (2015). Value-Based Reimbursement Decisions for Orphan Drugs: A Scoping Review and Decision Framework. PharmacoEconomics, 33. doi: 10.1007/s40273-014-0235-x

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist Review–a New Method of Systematic Review Designed for Complex Policy Interventions. Journal of Health Services Research & Policy, 10(suppl 1), 21-34.

Persson, U., Willis, M., & Ödegaard, K. (2010). A Case Study of Ex Ante, Value-Based Price and Reimbursement Decision-Making: Tlv and Rimonabant in Sweden. Eur J Health Econ, 11(2), 195-203. doi: 10.1007/s10198-009-0166-1

Plomer, A., Smith, I., & Martin-Clement, N. (1999). Rationing Policies on Access to in Vitro Fertilisation in the National Health Service, Uk. Reprod Health Matters, 7(14), 60-70.

Polikowski, M., & Santos-Eggimann, B. (2002). How Comprehensive Are the Basic Packages of Health Services? An International Comparison of Six Health Insurance Systems. J Health Serv Res Policy, 7(3), 133-142. doi: 10.1258/135581902760082436

Rauprich, O., Berns, E., & Vollmann, J. (2010). Who Should Pay for Assisted Reproductive Techniques? Answers from Patients, Professionals and the General Public in Germany. Human Reproduction. doi: 10.1093/humrep/deq056

Redmayne, S., & Klein, R. (1993). Rationing in Practice: The Case of in Vitro Fertilisation. Br Med J, 306(6891), 1521-1524.

Rocchi, A., Menon, D., Verma, S., & Miller, E. (2008). The Role of Economic Evidence in Canadian Oncology Reimbursement Decision-Making:

To Lambda and Beyond. Value Health, 11(4), 771-783.

Rooshenas, L., Owen-Smith, A., Hollingworth, W., Badrinath, P., Beynon, C., & Donovan, J. L. (2015). "I Won't Call It Rationing…": An Ethnographic Study of Healthcare Disinvestment in Theory and Practice. Soc. Sci. Med., 128, 273-281.

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. doi: 10.1016/j.healthpol.2010.09.008

Russell, J., Swinglehurst, D., & Greenhalgh, T. (2014). 'Cosmetic Boob Jobs' or Evidence-Based Breast Surgery: An Interpretive Policy Analysis of the Rationing of 'Low Value' Treatments in the English National Health Service. BMC Health Serv Res, 14, 413. doi: 10.1186/1472-6963-14-413

Schlander, M., & Beck, M. (2009). Expensive Drugs for Rare Disorders: To Treat or Not to Treat? The Case of Enzyme Replacement Therapy for Mucopolysaccharidosis Vi. Curr Med Res Opin, 25(5), 1285-1293.

Schnur, P. L., Hoehn, J. G., Ilstrup, D. M., Cahoy, M. J., & Chu, C.-P. (1991). Reduction Mammaplasty: Cosmetic or Reconstructive Procedure? Annals of Plastic Surgery, 27(3), 232-237.

Schreyögg, J., Stargardt, T., Velasco-Garrido, M., & Busse, R. (2005). Defining the "Health Benefit Basket" in Nine European Countries. The European Journal of Health Economics, 6(1), 2-10.

Shah, K. K., Tsuchiya, A., & Wailoo, A. J. (2014). Valuing Health at the End of Life: An Empirical Study of Public Preferences. Eur. J. Health Econ., 15(4), 389-399.

Shaw, L. M. A., Balen, A., Lenton, E., Brown, C., & Greenwood, B. (2002). National Health Service Provision for the Management of Infertility: The Case for Funding and Reorganization of Fertility Services in the Uk. Hum Fertil, 5(4), 167-174. doi: 10.1080/1464727022000199072

Simpson, L. A., & Cooper, J. (2009). Paying for Obesity: A Changing Landscape. Pediatrics, 123(Supplement 5), S301-S307. doi: 10.1542/peds.2008-2780I

Singer, P. A., Martin, D. K., Giacomini, M., & Purdy, L. (2000). Priority Setting for New Technologies in Medicine: Qualitative Case Study. Bmj, 321(7272), 1316-1318.

Specogna, M. B. (2010). Cessation Coverage in Argentina: A Qualitative Study About Its Barriers and Facilitating Factors. CVD Prev Contr, 5(1), 7-11. doi: 10.1016/j.cvdpc.2010.03.004

Stafinski, T., Mccabe, C., & Menon, D. (2014). Determining Social Values for Resource Allocation Decision-Making in Cancer Care: A Canadian Experiment. Journal of Cancer Policy, 2(3), 81-88. doi: 10.1016/j.jcpo.2014.07.002

Stevens, R. J. G., Stevens, S. G., & Rusby, J. E. (2015). The "Postcode Lottery" for the Surgical Correction of Gynaecomastia In nhs England. International Journal of Surgery, 22, 22-27. doi: http://dx.doi.org/10.1016/j.ijsu.2015.07.688

Stolk, E., Brouwer, W. B. F., & Busschbach, J. J. V. (2002). Rationalising Rationing: Economic and Other Considerations in the Debate About Funding of Viagra. Health Policy, 59(1), 53-63. doi: 10.1016/s0168-8510(01)00162-2

Tinghög, G., Carlsson, P., & Lyttkens, C. H. (2010). Individual Responsibility for What? - a Conceptual Framework for Exploring the Suitability of Private Financing in a Publicly Funded Health-Care System. Health Econ Policy Law, 5(2), 201-223. doi: 10.1017/s174413310999017x

Trosman, J. R., Weldon, C. B., Kelley, R. K., & Phillips, K. A. (2015). Challenges of Coverage Policy Development for Next-Generation Tumor Sequencing Panels: Experts and Payers Weigh In. JNCCN J Nat Compr Cancer Netw, 13(3), 311-318.

Trowman, R., Chung, H., Longson, C., Littlejohns, P., & Clark, P. (2011). The National Institute for Health and Clinical Excellence and Its Role in Assessing the Value of New Cancer Treatments in England and Wales. Clin Cancer Res, 17(15), 4930-4935. doi: 10.1158/1078-0432.ccr-10-2510

True, N. M. (2012). Removing the Constraints to Coverage of Gender-Confirming Healthcare by State Medicaid Programs. Iowa Law Rev., 97(4), 1329-1362.

Wichor M. Bramer, M. L. R., Jos Kleijnen, Oscar H. Franco. Optimal Database Combinations for Systematic Review Searches. submitted.

Wilkinson, S., & Williams, N. J. (2015). Should Uterus Transplants Be Publicly Funded? Journal of Medical Ethics. doi: 10.1136/medethics-2015-102999

Winquist, E., Bell, C. M., Clarke, J. T., Evans, G., Martin, J., Sabharwal, M., . . . Coyle, D. (2012). An Evaluation Framework for Funding Drugs for Rare Diseases. Value Health, 15(6), 982-986.

Wm Bramer, G. D. J., Ml Rethlefsen, F Mast, J Kleijnen.. A Systematic Approach to Searching: An Efficient and Complete Method to Develop Literature Searches. submitted.

Wong, A. M. (1995). The Inhumanity of Fairness: Rationing Resources for Reconstructive Breast Surgery. Cmaj, 152(4), 577-579.

Wong, G., Greenhalgh, T., & Pawson, R. (2010). Internet-Based Medical Education: A Realist Review of What Works, for Whom and in What Circumstances. BMC medical education, 10(1), 1.

Wong, G., Greenhalgh, T., Westhorp, G., Buckingham, J., & Pawson, R. (2013). Rameses Publication Standards: Realist Syntheses. BMC medicine, 11(1), 1.

Woolf, N. H., Burns, M. E., Bosworth, T. W., & Fiore, M. C. (2006). Purchasing Health Insurance Coverage for Smoking Cessation Treatment: Employers Describe the Most Influential Information in This Decision. Nicotine Tob Res, 8(6), 717-725. doi: 10.1080/14622200601004133

Wraight, W. M., Tay, S. K. L., Nduka, C., & Pereira, J. A. (2007). Bilateral Breast Reduction Surgery in England: A Postcode Lottery. Journal of Plastic, Reconstructive & Aesthetic Surgery, 60(9), 1039-1044. doi: http://dx.doi.org/10.1016/j.bjps.2007.03.002

Zelei, T., Molnár, M. J., Szegedi, M., & Kaló, Z. (2016). Systematic Review on the Evaluation Criteria of Orphan Medicines in Central and Eastern European Countries. Orphanet Journal of Rare Diseases, 11(1), 72. doi: 10.1186/s13023-016-0455-6

2 | Necessity argumentations across Western Europe

Published as: *Kleinhout-Vliek, T., de Bont, A., Boysen, M., Perleth, M., van der Veen, R., Zwaap, J., Boer, B. (2020). Around the Tables – Contextual Factors in Healthcare Coverage Decisions Across Western Europe. International Journal of Health Policy and Management, 9(9), 390-402.*

INTRODUCTION

Challenging coverage decision dossiers on a specific health care technology may show up on desks placed in different countries around the same time (Rosenberg-Yunger et al., 2011). Within their varying health care systems, decision makers are struggling with the exact same question: should our society pay for this? (Cerri et al., 2015). The processes to arrive at an answer to this question have not been aligned across Western Europe, though there are strong similarities. Western European systems do generally utilise some form of Health Technology Assessment for health care technologies (Franken et al., 2012; Kleijnen et al., 2012; Makady et al., 2017; Salas-Vega et al., 2016), which include technologies as wide-ranging as (orphan) drugs and medical devices, and "have clear objectives reflected in reimbursement criteria" (Franken et al., 2012). A formalised set of reimbursement criteria may, however, not necessarily wholly capture all reasons for or against coverage provided in a (publicly available) coverage decision document (Kleinhout-Vliek et al., 2017), which may result in differences between countries that might not be expected based on their respective formalised criteria sets (Maynou & Cairns, 2019; Nicod & Kanavos, 2012; Vuorenkoski et al., 2008). When surveyed, decision makers indeed frequently acknowledge the presence of additional elements that impact coverage decisions. Scholars group such additional elements under umbrellas like 'contextual factors' (Cerri et al., 2014; Cromwell et al., 2015; Csanádi et al., 2019; Hasman et al., 2008; Salas-Vega et al., 2016; Vuorenkoski et al., 2008; Williams et al., 2018; Wirtz et al., 2005).

What this umbrella of contextual factors holds exactly has recently attracted scholarly interest for 'decisions of value' in the health care field (Bærøe, 2018; Calnan, 2018; Peacock & Bentley, 2018; Williams et al., 2019). Williams et al. initiated the discussion with a literature review focused on meso-level decision making, exploring both 'inner' and 'outer' contexts and how these may affect these decisions of value (Williams et al., 2018). Many of the papers reviewed by Williams et al. give 'context' as a relatively abstract explanation, and as something that is not necessarily visible in the final decision document. Examples from this review include organisational/institutional forces (Eddama & Coast, 2008; Williams et al., 2008), political constraints (Miller et al., 2013), and economic climate or market forces (Bazzoli et al., 2007; Castro et al., 2014). The field of Science and Technology Studies (STS), and Asdal specifically, challenges us to not use contexts as what she terms non-specific 'explanatory resources' (Asdal, 2012; Asdal & Moser, 2012). Not content with doing away with context altogether either, Asdal encourages us to:

> [Grasp] the events of your study as, literally, unique events. (…) [Consequently,] the situation as the context that needs to be "recovered" is that which conditions or enables a specific utterance to happen. (Asdal, 2012, pp. 387-388)

In essence, she sees context as present in a specific decision situation, traceable to a specific moment in time: as *situation-specific*. This characteristic of situation-specificity, we argue, gives a handle on the contents of what Williams et al. term "greater levels of judgement and intuitions" on the part of

decision makers (Williams et al., 2018, p. 691). This chapter follows suit in conceptualising contexts expressly not as external explanations for a situation and not as enduring backdrops of whatever kind but rather as situation-specific entities that may be actively brought in by decision makers. Moreover, we will term contexts that are present in the final decision documentation, which provides a justification or rationale for the decision, 'contextual factors'. In this, we narrow down and specify Williams et al.'s original description of contextual factors (Bærøe, 2018). Of the contexts retrieved from literature by Williams et al., only two may be considered situation-specific, namely first, the information accessed by the decision makers, or its absence, and especially high levels of uncertainty regarding this information; and second, the presence of specific interests impacting the decision. That the technology and its information would vary per decision may be considered logical; the latter finding resonates with situation-specific contexts identified elsewhere for macro-level coverage decisions (Booth et al., 2007; Hasman et al., 2008; D. K. Martin et al., 2001b; Moes et al., 2017; Moreira, 2011).

Although some of the papers reviewed by Williams et al. describe minutely how the decision process unfolded, especially those included in the review by Vuorenkoski et al. (2008), little is known about how these contexts are integrated in the justification of, the rationale for, the decision specifically, as this often remains implicit in line with the surveys of policy makers described above. To be precise, no studies have, to our knowledge, compared such use of contextual factors across countries (comparative papers primarily focus on the use of Health Technology Assessment or related criteria such as cost-effectiveness and conclude that if decisions "vary", they do so due to "additional factors" (Clement et al., 2009; Nguyen-Kim et al., 2005; Nicod & Kanavos, 2012)). This chapter compares how situation-specific contextual factors are integrated in coverage decision documents in four Western European countries. Contextual factors will be operationalised here by means of a list of previously-described necessity argumentations (Kleinhout-Vliek et al., 2017). Necessity argumentations have been shown to be used in coverage decisions around the world but are not considered valid for every health care technology by everyone: their perceived validity varies per situation. These argumentations comprise a varied list and include e.g. the presence or absence of alternative treatments, the rule of rescue, the societal impact of (lack of) coverage, whether similar treatments are covered, medical necessity, and moral hazard considerations. Specifically, this research adds to existing literature by comparison of such contextual factor use in four widely varying health care technologies across four relatively similar countries in Western Europe, namely Belgium, England, Germany, and the Netherlands.

Research aim

This chapter will explore situation-specific contextual factors in health care coverage decisions. We will answer the following main research question: how do Belgium, England, Germany, and the Netherlands use contextual factors in health care coverage decisions generally, and how does contextual factors use compare across four specific decisions? This study is divided into two parts. Part 1 examines how contextual factors are used in decision documentations generally through

interviews and a workshop, enriched by document and web site analysis. Part 2 compares the contextual factors present in four specific decisions, Nivolumab (Opdivo®), benzodiazepines, smoking cessation therapy, and walking aids with wheels, taken in each of the four countries by examining the relevant decision documents.

METHODS

Approach

This study describes the outcomes of an international research collaboration of decision makers or policy advisors employed at four national health care institutes: National Institute for Health and Care Excellence (NICE) in England; the German Federal Joint Committee (*Gemeinsamer Bundesausschuss*, G-BA); the Health Care Knowledge Centre/National Institute for Health and Disability Insurance (HCKC/NIHDI) in Belgium; and the Dutch National Health Care Institute (*Zorginstituut Nederland*, ZIN) in the Netherlands with researchers from the Erasmus School of Health Policy & Management and the Erasmus School of Social and Behavioural Science, Rotterdam, the Netherlands on health care coverage decisions. For further information on the national health care institutes, please see section 'working procedures' in Appendix II. We provide an analysis of contextual factors in coverage decision documents generally (part 1 of the study) and in four specific cases (part 2) in these four countries.

We have chosen to compare countries in the understanding that comparative research need not be a linear-causal exercise, meaning here that a specific attribute of the country need not be offered as the explanation for an observed phenomenon (Deville et al., 2016; Krause, 2016). Rather, this method was chosen because careful comparison of arguments and statements is thought to result in clearer definitions of these and more succinct questions for future research to address (Deville et al., 2016). Specifically, we started with semi-structured, active interviews (Holstein & Gubrium, 2016) in groups. We enriched the interview data by document and web site analysis as well as an extensive member check (during the workshop, see below).

We have opted to further extend the data gathered in part 1 with case studies from the four countries (Creswell & Poth, 2017; Ragin, 2004). The reason for doing so is that comparison between cases that are similar (decisions on the same health care technology) taken in varying situations (the different countries) aids generating insight especially where valuation processes (cf. decisions of value (Williams et al., 2018)) are concerned (Lamont, 2012). In addition, the cases represent maximum variability (Lamont & Thévenot, 2000). In this way, by comparing the countries' general decision-making processes with four decisions specifically, we have chosen a small 'string of comparisons', which has the potential to bring strong clarification (Krause, 2016).

To operationalise these contextual factors, we will employ necessity argumentations as our sensitising concepts (Blumer, 1954). Necessity is an umbrella term that encapsulates disease severity and need-related argumentations, which are situation-specific because they vary strongly per

technology examined, and explicated and integrated by virtue of being used in decision outcomes. Necessity argumentations comprise patient-specific considerations, such as above-mentioned disease severity, medical necessity and need, but also dignity, human rights and impact on societal functioning, and whether the condition the patient suffers from may be construed as 'normal experience'. Related are the argumentations related to the patient population, such as whether the technology will be relevant for a small number of patients. Other considerations have to do with the technology and its availability, such as the presence or absence of alternative treatment and coverage of similar treatments, but also moral hazard (over-usage) considerations. Another type of argumentation is society-related and concerns argumentations like societal (or individual) responsibility and impact of (lack of) coverage on wider society.

Necessity argumentations are used to provide a justification of, a rationale for, coverage decisions in several European countries, including France, Germany, Sweden, the United Kingdom, and the Netherlands (Kleinhout-Vliek et al., 2017). These argumentations are largely considered valid not only by decision makers but also by the public. In the overall study, the Netherlands was taken as the entry point in terms of country selection and overall study focus, because necessity has been described and used as a criterion for Dutch coverage decisions (Hoedemaekers & Oortwijn, 2003; Stolk et al., 2001). England, Germany and Belgium were selected as comparator countries because of their similarities in terms of 'health care system objectives' such as equity and affordability but also transparency in decision making (Franken et al., 2012). This was expected to result in a publicly available decision, enabling our data analysis (Krause, 2016). The countries' specific working procedures for decision making may be found in Appendix II.

Data collection

The data collection consisted of two parts corresponding to the two parts of the study: first, a round of semi-structured group interviews with experts, general policy document and web site analysis and a workshop with all the authors (for an overview, see Table 1); and second, document analysis of the four decisions taken across the four countries (for an overview, see Table 2).

Part 1 of the data collection aimed to answer the first research question, was designed by authors AdB, BB, and TKV, and consisted of four semi-structured interviews, which were conducted in 2.5-3-hour conversations at each of the four health care institutes. These semi-structured interviews were conducted by BB, TKV, and JZ and held with one national expert, who acted as the primary point of contact per institute and subsequently agreed to co-author this chapter (MB, MP, and RvdV/JZ) and two colleagues, i.e. three interviewees per country, with the exception of Belgium, who opted out of authorship. These national experts hold the following roles: former Head of HCKC, Director of the Centre for Health Technology Evaluation at NICE, Head of the Department of Medical Advice at G-BA, and the Chair and Secretary of the Appraisal Committee at ZIN.

An exception was the Dutch group interview, where two experts agreed to co-author this chapter and only one additional colleague was present, and RvdV was interviewed one additional time by TKV (see Table 1).

Interview date	Description
24 February 2017	Appraisal committee ZIN focus group interview (RvdV, JZ, and a colleague)
17 March 2017	HCKC/NIHDI focus group interview (RM and two colleagues)
6 April 2017	G-BA focus group interview (MP and two colleagues)
23 June 2017	NICE/NHS-England focus group interview (MB and two colleagues)
18 October 2017	Additional interview (RvdV)
13 December 2017	Workshop at ZIN

Table 1 - Overview of interviews and workshops comprising Part 1 of the study

During the interviews, the interviewees all offered a deep understanding of their own country's health care decision processes and the use of contextual factors therein through giving a presentation on their decision-making procedure. This presentation was prepared beforehand, the interviewee giving the presentation was asked to provide information on the general coverage decision-making process in their country and its institutional bedding. They were also aware that the interviewers were interested in necessity argumentations in particular and thus provided information on these types of considerations as well. This was followed by a presentation by TKV on what the interviewers understood as contextual factors (that is, necessity argumentations) and an extended group interview on the role of these types of factors on the decision-making process generally. The questions used to guide these interviews may be found in Appendix I. The four semi-structured group interviews were audio-recorded, transcribed, and coded for necessity argumentation use. These analyses were enriched by analysis of relevant policy documents and websites as offered by the interviewees. This was generally achieved by perusing the general websites of the respective institutes and documents pertaining to their working procedures in case of lack of clarity.

The findings gathered during the group interviews (supplemented by the information from the web site and document analysis) were presented by TKV at a one-day workshop, held in the Netherlands, attended by all authors (13 December 2017). The explicit goal of this workshop was to gain deeper insight into the use of contextual factors in the different countries including similarities and differences as well as to provide a member check on the collected data. In addition, several analytical angles were discussed and the decision to add part 2 of the research was made. Part 2's aim was to examine the final documentation pertaining to four decisions taken across the four countries of interest. Three inclusion criteria for the case studies were formulated:

1. Necessity argumentations featured prominently in the final documentation of the coverage decision in the Netherlands;
2. The coverage decision outcome varies in the four countries, with at least one outcome deviating from the rest;
3. The cases together represent maximum variation in the patterns of use of the argumentation types (Kleinhout-Vliek et al., 2017).

For part 2, then, policy makers employed at ZIN (JZ and a colleague) aided with choosing eight cases that fulfilled the first criterion and that were sufficiently different in type of technology.

Inclusion criteria numbers 2 and 3 narrowed the list down to four cases: nivolumab (Opdivo®), benzodiazepines, smoking cessation therapy, and walking aids with wheels. Of the sixteen decisions (four countries times four health technologies), only fourteen yielded a document to analyse, and three decisions yielded two documents, to a total of seventeen documents. Approximately half were obtained through website searches of the relevant institutes, with the other half contributed by experts or colleagues working at that country's Health Care Institute (for an overview of the documents analysed, see Table 2).

Decision	Country	Document
Nivolumab	The Netherlands	Package advice nivolumab (Opdivo) including letter to the Minister of Health, Welfare, and Sports, dated 8 December 2015 (Zorginstituut Nederland, 2015)
	Belgium	Evaluation report day 90, Dossier 736 and 737, Second request report, dated 6 July 2016 (R.I.Z.I.V., 2016)
	England	NICE Technology appraisal guidance: Nivolumab for previously treated squamous non-small-cell lung cancer (TA483), dated 1 November 2017 (NICE, 2017a)
	Germany	File for Benefit Assessment in accordance with section 35a of SGB-V, Nivolumab (Opdivo®), dated 28 April 2016 (G-BA, 2016)
Benzodiazepines	The Netherlands	Package advice 2009, Publication number 274, dated 3 April 2009 (College voor Zorgverzekeringen, 2009a) Letter to the Minister of Health, Welfare, and Sports, Reimbursement benzodiazepines, dated 21 April 2008 (College voor Zorgverzekeringen, 2008b)
	Belgium	-
	England	Generalised Anxiety Disorder in adults, the NICE guideline on management in primary, secondary and community care (NICE) Generalised anxiety disorder and panic disorder in adults: management. Clinical guideline [CG113], dated January 2011 (NICE, 2011)
	Germany	-
Smoking cessation therapy	The Netherlands	Help with smoking cessation: insured care?, dated 30 June 2008 (College voor Zorgverzekeringen, 2008a) Smoking cessation programme: insured care! Publication number 276, dated 21 April 2009 (College voor Zorgverzekeringen, 2009b)
	Belgium	Effectiveness and cost-effectiveness of treatment for smoking cessation, HCKC, dated 2004 (van den Bruel, 2004)
	England	Stop smoking interventions and services, NICE guideline (NG92), dated 28 March 2018 (NICE, 2018)
	Germany	Regulation exclusion of medicines for heightening quality of life in accordance with Section 34 (1) Sentence 7 SGB V (Lifestyle Drugs), Annex II to Section F of the Medicine-Directive, dated 28 January 2017 (G-BA, 2017)
Walking aid with wheels	The Netherlands	Report Medical aids 2010, Publication number 286, dated 2 April 2010 (College voor Zorgverzekeringen, 2010)
	Belgium	Memorandum main working group number 2003/6.4, Main group 1.4, Walking Aids Adults, dated 14 July 2003 (Riziv-Inami, 2003)
	England	https://www.nhs.uk/conditions/social-care-and-support-guide/care-services-equipment-and-care-homes/walking-aids-wheelchairs-and-mobility-scooters/, dated 8 August 2018
	Germany	GKV-Spitzenverband Update of the product group 10 "Walking aids" of the aid directory according to Section 139 SGB- VvomAf, dated 27 August 2018 (GKV-Spitzenverband, 2018)

Table 2 - Overview of documents analysed comprising part 2 of the study. Note: for the English Benzodiazepines decision, please note that the second document analysed (Clinical guideline 113) is based on the former document.

TKV was aided by a student in analysing all coverage decision documents for the primary arguments separately, again using the list of necessity argumentations (Table 3) as codes, using good language skills in Dutch (for the Dutch and Belgian documents) and English and sufficient aptitude in reading German, aided by online translation and where needed by professionals who were directly involved in the coverage decision processes. The coding was done by TKV by specifically searching for forms of these twenty necessity argumentations. The outcomes were discussed by AdB, BB, and TKV in several iterative meetings, and all other collaborators commented on the resulting analysis and agreed upon the final text.

RESULTS

This section provides an overview of how contextual factors, operationalised as necessity argumentations, generally play a role in each country, followed by a section on the use of contextual factors in the four case studies. A short introduction to the decision process in the different countries (derived from the website and policy document analysis and the presentation given by the interviewees during the semi-structured interviews) may be found in Appendix II.

Part 1: Use of contextual factors in general

From the presentations during the semi-structured group interviews in each of the four countries, as well as from the policy document and web site analysis, we conclude the following after introducing the interviewees to the list of necessity argumentations, sensitising them to the topics.

According to the interviewees, their respective national health care institutes all use necessity argumentations as contextual factors in their decisions in addition to the formalised criteria (outlined below). Members of the research team easily provided examples where, in their eyes, a wide variety of these non-formalised, situation-specific contextual factors were employed. These examples included but were not limited to: inclusion of smoking cessation (Belgium, see case study below); exclusion of homeopathy (England, decided by the National Health Service (NHS), perceived by interviewee as a necessity argumentation); exclusion of over-the-counter medicines (England, due to low cost per patient); positive "Nikolaus decision" (Germany, due to a small number of patients); exclusion of immunisation for travel in free time (Germany, this falling under personal responsibility). Notably, the workshop attendees underlined that not all might be identifiable in the final decision: sometimes these decisions were, apparently, 'gut decisions', with reasons remaining implicit.

More generally, reflecting on the list of contextual factors, the Belgian interviewee stated:

> [Contextualisation] will remain discursive all the time. (workshop, 13.12.2017)

This observation highlighted that contextualisation is left open to be established in argumentations in a deliberative setting: 'around the table'. It also accentuated the difference between contextual factors, which are situation-specific, and formalised criteria, which are not. This does not mean that of all the necessity argumentations, none have been formalised into criteria. In fact, the national health care institutes have all formalised the use of these contextual factors but to different extents.

Least formalised is Belgium, where no list of stringent criteria exists: rather, decisions are made and contextual factors formulated by different working groups. Some necessity argumentations are explicitly reflected in the criteria used in the Unmet Needs Programme, which features therapeutic need (discomfort, life expectancy, quality of life) and societal need (budget impact, incidence/prevalence) but this list is not widely used for decision making (Cleemput et al., 2018).

In England, the situation could be called most intricate. NICE general procedure turns on clinical and cost-effectiveness, with thresholds for opportunity costs. Disease severity is formally included in the application of the 'end-of-life' criteria, which allows additional flexibility only where a patient's life expectancy is less than two years. It is also used in the deliberative decision-making process where the independent committee has the flexibility to accept cost-effectiveness estimates that are higher than what would otherwise be considered value for money. Moreover, the NICE pre-selection procedure employs elements like chronicity and acute death as criteria.

For Germany, 'necessity' is one of three formalised criteria described in the German Social Code, Book 5 (SGB-V). However, it has not been specified to a great extent; the criteria in Germany are considered "generally formulated" (Patera & Wild, 2014). Disease severity does influence the level of evidence required in G-BA's decisions (varying from anecdotal evidence to randomised controlled trials).

The country with the highest formalisation is the Netherlands; ZIN operationalises necessity as one of four formal package criteria (next to effectiveness, cost-effectiveness, and feasibility). This is explicated in their formal documentations as disease severity, or individual burden of disease, and individual cost considerations (Zorginstituut Nederland, 2013, 2017).

Part 2: Use of contextual factors in four case studies

The four decisions to be examined across the four countries, namely nivolumab, benzodiazepines, smoking cessation therapies, and walking aids with wheels, were selected per the criteria mentioned above, the documents analysed per decision are available in Table 2. Table 3 gives an overview of the used contextual factors per decision document per country.

Case study 1: Nivolumab

All four countries decided on reimbursement of nivolumab for treatment of previously-treated squamous non-small-cell lung cancer (NSCLC).

The expectation of a large budget impact necessitated placing nivolumab in the Dutch package lock, awaiting an advised decision by ZIN and, potentially, subsequent price negotiations with the manufacturer (Schippers, 2015). The Institute did indeed advise the Minister of Health against

Contextual factor	Description	Nivolumab	Benzodiazepines	Smoking cessation therapies	Walking aids with wheels
Definition of Illness	Whether the ailment is considered an illness for which treatment is necessary	NL		NL; BE; EN; DE	
Equity/ Fairness/ Justice	Whether coverage would be necessary to counter injustice/inequity/lack of fairness in (access to) treatment			NL; EN	
Individual Cost	Whether lack of coverage would stop patients from buying necessary care themselves due to prohibitive cost	NL	NL	NL	NL
Individual Responsibility	Whether the individual is considered responsible for paying for this treatment			NL	NL
Medical Necessity	Whether or not a treatment is considered to be "medically necessary" or a "medical necessity"		NL		
Morbidity/ Severity	Whether the physical and/or psychosocial morbidity associated with a certain ailment constitutes such a need that coverage is considered necessary	NL; BE; EN; DE	NL; EN	NL; BE; EN	BE; EN; DE
Need	The extent to which the patient is considered to be in need for which treatment is necessary	NL; BE; EN; DE			BE
(No) Alternative	Whether or not viable alternatives are considered to be present which would make coverage more or less necessary	NL; BE; EN	NL; EN		DE
Patient-Diagnosis	Whether an illness is self-reported rather than diagnosed by a doctor		NL; EN		
Range of Normality	Whether the experience of the patient is considered normal or abnormal to such an extent that coverage is deemed necessary				NL
Similar Treatments	Whether similar treatments are covered or not (meaning that this *type* of treatment is considered necessary)			NL; DE	
Societal Impact	Whether coverage is considered necessary to allay the impact this disease has on people beyond the patient			NL; BE	
Societal Functioning	Whether coverage would aid a person's necessary functioning in society	BE; EN	EN	NL; DE	BE; EN; DE
Vulnerability/ Compassion	Whether a compassionate response to vulnerable groups, e.g. children, in the form of coverage is considered to be a necessity			NL; BE	
Substitution	Whether other (e.g., heavier dosage or more expensive than necessary) medicines or care would be consumed or used by patients as a result of a negative coverage decision		NL		
Under-consumption	Whether less medicines or treatments than necessary would be consumed or used by patients as a result of a negative coverage decision (the opposite of 'Moral Hazard')			NL	

Table 3 - Overview of contextual factors operationalised as necessity argumentations and their respective descriptions (Kleinhout-Vliek et al., 2017, 2020) combined with information on which factor was used in which decision in which country. The argumentation types that were not present in this data set (namely, Dignity, Human Right, Moral Hazard, Rule of Rescue, Small Number of Patients, and Societal Responsibility) were omitted from this table.

including nivolumab in the basic benefits package unless a price reduction of at least 40% could be achieved: the treatment was considered effective but not cost-effective enough. In their decision documentation, the Institute argued that: "the burden of disease is high" as it is considered "a non-curative disease" with "limited life expectancy". Interestingly, an alternative treatment (docetaxel) is mentioned in the pharmaco-economic report). Moreover, "the costs for the treatment is so high patients cannot be expected to pay" (Zorginstituut Nederland, 2015). The Minister did negotiate and although the final price has not been made public, nivolumab is now available on the Dutch benefits package.

For Belgium, the official documentation retrieved is the 'day 90' report from the Committee for 'advance compensation' for Pharmaceuticals and specifically the 'second request' document, which gives the latest insight into motivation concerning reimbursement and the details of the final coverage arrangement for nivolumab. The necessity argumentations visible are "weakness and fatigue, coughing, shortness of breath, pain" and a median survival number. Moreover, considerations around the quality of life feature in the comparison with doxacetel (R.I.Z.I.V., 2016). This resulted in a positive decision.

As for ZIN, NICE's Technology Appraisal Committee considered nivolumab effective but not cost-effective enough. Moreover, the uncertainty of the evidence was deemed too great. For these reasons, additional characteristics were considered: in the social value judgement several necessity argumentations were brought to the fore. The committee noted a high need due to lack of alternatives, with a high morbidity, a low life expectancy, and "symptoms which are difficult to manage" (NICE, 2017a). This resulted in a recommendation of funding but *only* through the Cancer Drugs Fund, which happened after a renegotiation on the price of nivolumab with manufacturer Bristol-Myers Squibb (NICE, 2017b).

In Germany, drugs are covered with market entry, followed by a decision on additional benefit by G-BA, which may be used in turn for price negotiations. In the case of nivolumab, the Institute for Quality and Science in Health Care (in German: *Institut für Qualität und Wirtschaftlichkeit im Gesundheidswesen*, IQWiG) provided evidence concerning mortality, morbidity, quality of life, and several adverse event categories, noting considerable added benefit on all counts for NSCLC patients on nivolumab to benefit the decision made by G-BA. In terms of necessity argumentations, the report mentions the severity of the disease, the low absolute 10-year survival rate, and the "need for drugs". The Individual Cost is mentioned but not clearly as an argumentation (G-BA, 2016). This was sufficient grounds not to exclude nivolumab from coverage.

Case study 2: Benzodiazepines

In the Netherlands, ZIN advised the Minister of Health against continued coverage of benzodiazepines (College voor Zorgverzekeringen, 2008b, 2009a). The primary reasoning was that although the indication was for short-term use, prescriptions for benzodiazepines often ended up to be for chronic use. The Institute recommended improving the guidelines, but consultation with field parties revealed difficulties in defining eligibility criteria. Therefore, even despite the disease severity

mentioned and the fact that short use of benzodiazepines following the applicable guidelines may be medically necessary (but chronic use not), the large amount of chronic users, the possibility that denying coverage would lead to patients choosing another, covered, medicine and the fact that the cost for one episode of benzodiazepines is approximately €12-16 were decisive in the negative coverage decision. The exception was made for three clearly defined indications: epilepsy, anxiety disorders, and multiple psychiatric disorders.

For the Belgian case, treatment of benzodiazepines is not covered by the health insurance. NIHDI has never appraised benzodiazepines as they have never received a request to do so.

In England, NICE guidelines recommend short-term use of benzodiazepines for several medical indications, which have separate appraisal documents. We analysed the case of adult Generalised Anxiety Disorders (GADs), which highlights several risks for GPs prescribing benzodiazepines, which fall under necessity argumentations. First, emphasis is laid on the high frequency of over-use due to tolerance and dependence. The documentation does highlight that use is only recommended in case of non-response to other medicines, which can be classed as an argumentation in favour of coverage. People with GAD are described to have "long-standing and often uncontrollable worries and negative thoughts" which affect their "many areas of their lives, particularly relationships, self-esteem, daily activities, employment, work life and education" (NICE, 2011).

G-BA has, like NIHDI, never received an appraisal request for benzodiazepines, but in contrast to Belgium, benzodiazepines are covered by the German sickness funds. The reason is that in Germany all drugs are covered after market entry except when they are specifically excluded either by law in the SGB-V or by G-BA. A systematic assessment of additional benefit (for the sake of price negotiations) was only established in 2011 by the Act on the Reform of the Market for Medicinal Products. Benzodiazepines were on the German market in 2010 already and are thus part of a "historical benefit package."

Case study 3: Smoking cessation therapies

This analysis is of decisions for both psychotherapeutic and pharmaceutical smoking cessation therapies.

In the Netherlands, ZIN dealt with several iterations of the coverage decision for smoking cessation therapies: they were covered, then no longer covered, and then covered again. One reason is that it became a political issue with changes in government affecting the reimbursement status. The final advice to the Minister has been to cover stop advice, intensive forms of interventions for behavioural change, and nortriptyline but not to several other interventions. The final decision hinged, at least partially, on the effectiveness: in essence, the health gain even a few extra successful stop attempts would yield, also visible in the argumentation that "a smoking cessation programme can reduce the damage caused by smoking to others". In this extensive decision-making process, several necessity argumentations played a role (College voor Zorgverzekeringen, 2008a, 2009b). First, the individual is personally responsible for his health, and the costs for the individual were low, reducing the perceived necessity of coverage. Even so, there were indications that more people

would attempt to quit smoking if it was covered, and especially those with a lower socio-economic status. Moreover, smoking was defined as an addiction and bad for health, leading to quality of life loss, causing damage to others, including unborn children and infants, treatment for which is usually covered by the basic benefits package.

Smoking cessation therapies have never formally been discussed by the committee for Reimbursement of Medicines (Tegemoetkoming Geneesmiddelen) at Belgian NIHDI. Not the mutualities but the Belgian National Cancer Initiative covers €20 of the first eight sessions of smoking cessation therapies (Kankerplan, 2008-2010) and from 1 January 2017, smoking cessation aid by a tobaccologist is reimbursed in Flanders (overheid, 2019). As in the Netherlands though, these therapies have not always been covered for everyone. The compulsory health insurance covered smoking cessation therapies for all pregnant women, revealing a potential Societal Impact argumentation, though this was rarely used. In 2004, HCKC published a report that all smoking cessation therapies are cost-effective that informed what is now a primary part of the National Cancer Initiative. This report was not a coverage decision, but analysis yielded that smoking was defined as a habit, strong descriptions of the morbidity caused by smoking, including the increased risk of having a baby with low birth weight (van den Bruel, 2004).

In England, smoking cessation therapies are available to all citizens over 12 years of age, for which NICE most recently published guidelines in 2018. These highlighted that "smoking is the main cause of preventable illness and premature death" and showed that all smoking cessation therapies were cost-effective enough to stay below the QALY threshold. The guidelines underscore that there are social inequalities in terms of tobacco use, which "make a significant contribution to inequalities in health" (NICE, 2018).

Germany is the only of the four countries analysed in this project that does not cover smoking cessation therapies, as they are excluded from the benefits package by law (SGB-V §34). The title of the analysed documentation reveals they are considered "lifestyle medicines" aimed at improving quality of life, but no additional contextual factors are visible in the documentation.

Case study 4: Walking aids with wheels (rollators)

We analyse walking aids with two, three or four wheels, also called rollators, for adults.

In the Netherlands, the Minister of Health was advised by ZIN to no longer include mobility aids in the basic benefits package, and two necessity argumentations were given for this. First, the cost for walking aids with wheels, crutches, and many more walking aids was considered too low and secondly, "a walker with wheels is for common use", meaning that it was, in Dutch society, considered normal to need one at a certain age (College voor Zorgverzekeringen, 2010).

The Belgian NIHDI has published prescription guidelines which cover walking aids with wheels for those unable to stand up independently or safely (Riziv-Inami, 2003).

In England, this intervention has not been considered by NICE but the NHS loans walking aids with wheels for free, after assessment by a physician. No specific appraisal documents were available for walkers with wheels, as they are only discussed in guidelines for specific conditions, which

only specify that "you or someone you know" needs to have "difficulty walking or getting around (mobility)" (NHS, 2018).

Despite the fact that G-BA has not appraised walking aids with wheels specifically, they are available to all German citizens from their insurers upon indication: anyone experiencing reduced physical mobility or loss of balance is considered eligible, as long as other mobility aids have proved insufficient. This list is managed by the National Association of Statutory Health Insurance Funds (*GKV-Spitzenverband*) (GKV-Spitzenverband, 2018).

Comparison of contextual factor use in the four cases

To start, across these cases, similar patterns in contextual factor use lead to similar decision outcomes. This is visible in five instances. First, there are strong similarities between the decisions for nivolumab in Belgium and Germany, which both strongly rely on the severity of the disease and patients' quality of life and need, as well as the low survival rate. The other two decisions for nivolumab are highly reminiscent of these argumentations, but the Dutch and English both used them as input for price negotiations (which means the initial decision was negative). England and the Netherlands also strongly overlapped in their argumentation pattern for benzodiazepines, invoking the large amount of chronic users or over-usage as a reason to limit coverage, while acknowledging the difficulties experienced by the patient as well as the lack of alternatives. Fourth, there are some similarities between the Dutch and English decision for smoking cessation therapies, too, though the Dutch decision-making process featured a far greater number of contextual factors. Finally, the Belgian, English, and German decisions for walking aids with wheels also show strong overlap: if you are unable to stand up safely in these countries, a walking aid with wheels is provided for in some way.

Moreover, it would seem that the number of contextual factor types considered valid in one or more of the cases varies between the countries: the Belgian decisions feature seven types, the English also seven types, the German decisions six types, compared to sixteen types of contextual factors present in the Netherlands. This shows that the Dutch decisions generally feature a high amount of contextual factors in this data set.

Finally, half the (final) decisions were not taken by the institutes analysed. This pertains to the Belgian and German benzodiazepines decisions, the Belgian and German smoking cessation therapies decisions, the English and German walking aids with wheels decisions and, to an extent, the Dutch nivolumab decision as it differed from the initial ZIN advice, and the English nivolumab decision as it was covered through the CDF rather than a positive decision by NICE. Many of these did, however, yield a document or web site to analyse (only two benzodiazepines decisions lacked such a decision document).

DISCUSSION

Contextualisation has recently been a raised as an important topic of interest for both policy practice and research regarding health care 'decisions of value' (Williams et al., 2018). Our cross-country research team has defined contextual factors as *situation-specific* considerations, following Asdal (2012). This enabled us to examine the coverage decision processes in four Western-European countries: Belgium, England, Germany, and the Netherlands, and specifically, to establish where contextual factors are used, both generally (part 1 of the study) and in four decisions taken across the four countries specifically (part 2). We have operationalised these situation-specific contextual factors using a list of previously-described necessity argumentations that are used across Europe and generally vary per decision (Kleinhout-Vliek et al., 2017).

We draw the following conclusions. From part 1 of the study, we conclude that situation-specific contextual factors, operationalised as necessity argumentations, are present in decision-making processes at the HCKC/NIHDI in Belgium; NICE in England; the German G-BA; and the Dutch ZIN in the Netherlands. Some necessity argumentations have been formalised into criteria to be used for every decision in theory (individual burden of disease in England, Germany, and the Netherlands, and individual cost considerations in the Netherlands) (Commissie-Dunning, 1991; Franken et al., 2015; Stolk et al., 2002; Zorginstituut Nederland, 2013). This may raise the question whether these would fall outside the definition of contextual factors as situation-specific – however, our data show that not every decision analysed uses the formalised contextual factor(s) in practice. Specifically, expert interviewees underline that these contextual factors are determined in deliberation, 'around the table': a setting present in some form at every institute studied (Patera & Wild, 2014). They offered many examples of specific decisions in which such factors were considered, though underlining that not all of these factors had been explicated in their respective documentations.

From part 2, we conclude that similar patterns in contextual factor usage lead to similar decisions in the countries studied (Belgium and Germany for nivolumab; England and the Netherlands for nivolumab and benzodiazepines; and Belgium, England, and Germany for walking aids with wheels). This is an important conclusion, which may serve to encourage exchange between decision makers in different countries on more qualitative aspects of coverage decisions in addition to the current collaborations on the more quantitative aspects. However, the decisions are still sufficiently different (in fact, they were explicitly selected as having different outcomes across the four countries) to preclude much more than exchange. The use of contextual factors in decisions, we would argue, would need to remain at the discretion of local decision makers. In this data set, Dutch decisions employ the widest variety of contextual factors, and most often employs an argumentation type no other country employs. This with the notable absence of the societal functioning of the patient, which is a common consideration in Belgium, England, and Germany. This shows that in a sense, the argumentations are not 100% situation-specific. Instead, they are part of typifications and these types are identifiable across situations but not consistently. It is thus the *pattern* of contextual

factors that is truly situation-specific (D. Martin et al., 2001a; Singer et al., 2000). Future research could further address this idea of patterning and how it influences coverage decisions. Finally, for all four institutes some decisions in this data set are taken or retaken by another actor, which often sometimes means that the documentation for the final decision is not publicly available.

Taking a step further in conceptualising contextual factors, we conclude that they are explicated around the table, in deliberation (Kleinhout-Vliek et al., 2020). This establishing around the table is the first element: we also conclude that contextual factors need to be *actively integrated* in the decision documentation, as not all factors established in deliberation seem to be present in the document. Many scholars have in fact described, often based on interviews, that many such contextual factors remain implicit, either in both the deliberation and the documentation or in the documentation only. This is perhaps to be expected, as localised processes of meaning-giving by the decision makers themselves are described as implicit (Hughes & Light, 2002; Rooshenas et al., 2015). We conclude, as Mann puts it, that contextualisation is an intervention (Mann, 2015) and as such an active, situated process. As a consequence, this conceptualisation of contextual factors attunes us to effect that decision makers have on how decisions are justified and the expertise they bring to bear therein (Kleinhout-Vliek et al., 2020). The fact that these differences in explication of factors exist is of particular interest as it provides more insight into the how and why of evidence-informed deliberative processes (Baltussen et al., 2016; Hall, 2017). Further research should address why and how some factors remain implicit, whereas others are not only explicated but actively integrated in the final text.

Many of the documentations that should provide a justification/rationale for the final decision were, in fact, absent and an alternative document was analysed if present. For some, we analysed the pre-final decision (nivolumab for England and the Netherlands, both negative decisions), in other cases, the decision was not taken by the national health care institute (walking aids with wheels for England and Germany), in others again, no decision was visible at all. This underscores that health care coverage decision making is a process that involves many people in many places. It is a question of definition ('does a decision on technology X fall under the remit of our national health care institute?') but also a question of agenda setting (e.g., NIHDI and G-BA had never received a request to appraise benzodiazepines). This backdrop to the decision is something that is not explicitly integrated into the documentation but is definitely actively shaping the final decision.

This high prevalence of decisions for which the final documentation is not available is a particularly salient finding, as this highlights that the transparency of some decisions may be limited. Because these (final) decisions are made in another setting, the argumentations underlying the decision remain unknown. This is intriguing as many of the institutions studied are seeking to make processes more transparent in pursuit of increasing the legitimacy of their decisions (Franken et al., 2012). A more transparent decision is considered to heighten the legitimacy of this type of public decisions in general (Alonso-Coello et al., 2016; Daniels, 2000; Daniels et al., 2016; Daniels & Sabin, 1997, 1998, 2008; Guyatt, Oxman, Kunz, et al., 2008; Guyatt, Oxman, Vist, et al., 2008). Moreover, many authors hold that the coverage decision process should be based on consistently-applied,

formalised criteria, visible in the documentation pertaining to the decision and that having highly formalised criteria would potentially enable more rational, better-justifiable decisions (Bærøe & Baltussen, 2014; Baltussen & Niessen, 2006; Giacomini et al., 2000; Jansen et al., 2017; Kapiriri et al., 2009). Yet, the Dutch nivolumab decision in particular demonstrates that having highly formalised criteria, as is the case in the Netherlands, does not preclude this particular type of decision making. Further research will need to carefully consider the ways in which these 'invisible' decisions enable and restrain deliberative coverage decision-making practice and how this relates to the legitimacy of these decisions (de Fine Licht, 2011).

Strengths and limitations

This chapter covers a vital topic as it successfully visualises the contextual factors employed in coverage decisions generally and four decisions specifically in four countries. This chapter does so without resorting to using the general health care system characteristics as an explanation but instead seeks to draw more specific conclusions. Another strength of the chapter lies in the methodology and especially the case-study selection, which was both grounded empirically through expert interviews and enriched by theoretical interest through the formulated criteria.

As the Netherlands was taken as the entry point for this study, both in terms of content as in terms of the place of work and residence of the majority of the authors, this will have had a major impact. In particular, this is visible first, in terms of the case selection (as Dutch interviewees were asked for the initial list of potential cases); second, in the level of detail acquired in the case descriptions, which is relatively high for the Dutch decisions. It is also likely that it may have affected the research question itself. Asking questions concerning formalisation, or operationalisation, of criteria may be considered a typically *Dutch* pursuit, as visualised by the fact that the Netherlands has a rich history in terms of seeking to explicate decision criteria in this context (Commissie-Dunning, 1991; Stolk et al., 2002; Stolk et al., 2001).

The use of necessity argumentations has narrowed the subject down content-wise to considerations that are likely to be present in decision documents. Future research could investigate how what remains implicit impacts decision making. It is also important to note that the decisions studied span the last fifteen years, and it is likely that considerations are weighed or valued differently across that time span. Future research could address how the use of contextual factors may change over time.

As Bærøe noted (Bærøe, 2018), there is a difference between approaches aiming to formulate a comprehensive list and those that would hold that this is impossible. Although we *have* chosen a limited list as it facilitates this research, we do not necessarily believe that an exhaustive list would be possible. Normatively this may well prove problematic; for this study, we are concerned with describing what is rather than with what ought to.

CONCLUSION

This study aims to be part of answering the recent call for research in aid of understanding of practices of contextualisation. As health care coverage decisions are a particularly fruitful area to study these practices in, we have compared the use of contextual factors, defined as situation-specific considerations, in documentations that provide a justification or rationale of these decisions as offered by HCKC/NIHDI (Belgium), NICE (England), G-BA (Germany), and ZIN (the Netherlands). To study these, we build on group interviews with three national experts per institute, document and web site analysis, and a workshop with one to two experts per country (together part 1), and the analysis of four different case studies across these four countries, which varied greatly in terms of type of technology (part 2). From this data set, we conclude that these four national health care institutes all utilise situation-specific contextual factors in their decision documents. These contextual factors are employed 'around the table', that is, established in deliberation. Though some may remain implicit, others are not only explicated but *actively integrated* in the decision documentation, thus strengthening the decision by making it more sensitive to the case at hand. Moreover, in this data set, there are strong similarities in terms of how these contextual factors are used: similar patterns of contextual factor use lead to similar decisions in different countries. These observations do not use context as non-specific explanatory resources, as critiqued by Asdal in particular (Asdal, 2012; Asdal & Moser, 2012) but instead focus on the people and their processes required to actively integrate these considerations. It also calls for future research on patterning of these contextual factors in deliberative settings. Not all decisions are taken 'around the table', however. Half the decisions were taken or retaken in another setting, with the documentation to back up the final decision sometimes completely absent. We note this may impact the legitimacy of these decisions and call for future research efforts on how this may affect the daily practice of coverage decision making.

APPENDIX I – INTERVIEW GUIDE FOR SEMI-STRUCTURED INTERVIEWS

- Please tell us about yourself, your institute's coverage decision-making process, its institutional setting, etc. (15 min. presentation by interviewees, prepared beforehand)
- What is your general impression of how necessity and its related notions (e.g. need and situation of the patient and the responsibility of society) are used in your country? Could you give one or two examples of a decision where necessity was important and/or decisive?
- In your opinion, is there anything that could be improved in decision-making processes based on necessity in your country? Could you give an example of a decision for which this would have made a difference?
- What are your thoughts on the finding that these necessity argumentations may vary in usage and validity? Is there anything that resonates with your experience particularly? Could you give examples?
- To what degree are these findings understandable and relevant for your work? How would they impact future decisions?
- What (type of) research outcomes would be most valuable for your work in terms of this study? How do you think they might be used at your workplace?
- What it has been like to participate in this interview, what has been good, and what could be improved upon?
- Is there anything else you would like to share with us before we conclude the interview?

APPENDIX II - WORKING PROCEDURES PER COUNTRY

The Netherlands

The Dutch health care system is insurance-based: all Dutch citizens are obliged to take out health insurance with a private insurance company. However, the bare minimum for which they are ensured, the basic benefits package, is the same for all citizens and is established by the government. The Dutch National Health Care Institute (ZIN) advises the Minister of Health in his or her decision as to what should, and should not, be covered (Couwenbergh et al., 2013). The mission of the Institute is to provide "no less than needed but no more than necessary" health care (as shown on their website, accessed 17 January 2019), a mission that is worked out through advice on the basic benefits package but also, for example, setting of quality standards.

ZIN employs four *a priori* established criteria, also called "package principles": effectiveness and cost-effectiveness of the technology, feasibility of coverage (including budget impact), and necessity, which includes individual severity of illness and necessity of coverage. The process at the Institute is guided by Institute employees, who also write the final advice document and all previous drafts. The process starts with a scoping session with stakeholders, like patient organisations, health care providers, pharmaceutical companies, and health insurers, who are invited to comment on the process. The scoping is followed by an extensive examination of the scientific evidence in the Scientific Advisory Council. The third step is the appraisal, or contextualisation, of this evidence in the Advisory Committee Package in a setting that is open to the public. Generally, some of the stakeholders will be present here as well, and there is an opportunity to contribute to the deliberations, which ought to consider wider societal values pertaining to the coverage. The final advice is then formulated and approved by the Board of Directors, before it is sent to the Minister (Zorginstituut Nederland, 2016, 2017).

Belgium

Belgium, similar to the Netherlands, operates on a social health insurance system basis, which mandates citizens to take out basic health insurance with an insurer. These insurance organisations, called mutualities, operate on a not-for-profit basis. Decisions on what is covered by this compulsory health insurance are discussed at the National Institute for Health and Disability Insurance (NIHDI). NIHDI allocates the national health care budget, sets standard prices for treatments, and inspects the mutualities, supervised by the Minister of Public Health and Social Affairs. The Health Care Knowledge Centre (HCKC), in contrast, is responsible for research and scientific advice concerning the basic health insurance, taking on a consultancy role for policy makers, who may commission such research (Centre, 2017). This research generally takes a multi-disciplinary approach. HCKC counts over 45 experts with a background in medicine, economy, sociology, law, and/or ethics. They formulate recommendations together, aided where necessary by external parties and in collaboration with the Board of Directors (Centre, 2019). The coverage decision proposals of NIHDI are sent to the Minister who makes the final decision.

England

The NHS, the National Health System in the United Kingdom, is publicly funded, free at the point of delivery, based on clinical need, not the ability to pay, and is aimed at meeting the needs of everyone (Grosios et al., 2010). The National Institute for Health and Care Excellence (NICE) provides guidance and advice concerning health care services in England. NICE was established in 1999 in an attempt to tackle the so-called postcode lottery in access to new technologies. The Health and Social Care Act of 2012 notes that developing its advice or guidance, NICE must have regard to the broad balance between the benefits and costs of the provision of health services or of social care in England, the degree of need of persons for health services or social care in England, and the desirability of promoting innovation in the provision of health services or of social care in England. The NICE single technology appraisal is the primary, standardised way NICE evaluates old and new forms of care but also multiple technology appraisals and fast track appraisals are available. Moreover, NICE's guideline programme considers the care and services suitable for most people with a specific condition or need and people in particular circumstances or settings (NICE, 2016).

The starting point for technology appraisals is the evidence submitted by the company, where when it comes to a guideline, NICE uses academic collaborating centres to bring together the evidence. In technology appraisals, this evidence is then considered by the ERG and added to it evidence from external parties, which may include clinical specialists, commissioning experts and patient experts. The appraisal committee subsequently convenes to appraise the available evidence in terms of clinical and cost-effectiveness, consider advice given by NICE, board, and drawing on social value judgements, including those informed by the Citizens Council, to come to a first Appraisal Consultation Document (ACD) containing preliminary recommendations (Dakin et al., 2015). This Citizens Council, which NICE has had for many years, helps frame the social value judgements that the independent committees are asked to consider when formulating NICE guidance. Once it is published, all stakeholders, including the general public, may comment on the ACD, which are potentially taken along in NICE's final recommendation (NICE, 2013).

Germany

The health care system in Germany is social insurance based, like the Netherlands and Belgium, with the statutory health funds (SHFs) covering circa 90% of the German population, and private insurance covering the rest. The costs for insurance are shared between employers and employees, and German citizens are free to choose among the SHFs (ISPOR, 2009). All SHFs are represented in the National Association of Statutory Health Insurance Funds (GKV-Spitzenverband), the SHF umbrella organisation. The Gemeinsamer Bundesausschuss (G-BA) is the highest, independent organisation in the German health care system, playing a major role in the decision-making process concerning which services are covered by the SHFs and issuing directives concerning the national benefits package. It brings together four major organisations representing physicians, dentists, hospitals, and insurance funds respectively. The German Social Code, Book 5 (SGB-V) sets out the

lawful responsibilities of G-BA, specifying rules for reaching agreement, appointing members and involving patients and third parties.

The main decision-making body of G-BA, the plenum, is a deliberative setting in which decisions are reached concerning which types of care are in- or excluded. The plenum comprises thirteen voting members and five patient representatives. Of these thirteen voting members, three are impartial (including the chair), five are representatives from the GKV-Spitzenverband, and the other five are care provider representatives. Cases are prepared by one of the nine subcommittees for discussion in the plenum, which have their own expertise (Gemeinsamer Bundesausschuss). Cost-effectiveness data in particular are prepared by IQWiG, the independent federal organisation set up to evaluate medical efficiency, quality and effectiveness of treatments. These data are considered in the plenum next to consultations by experts and practitioners.

REFERENCES

Alonso-Coello, P., Schünemann, H. J., Moberg, J., Brignardello-Petersen, R., Akl, E. A., Davoli, M., . . . Rosenbaum, S. (2016). Grade Evidence to Decision (Etd) Frameworks: A Systematic and Transparent Approach to Making Well Informed Healthcare Choices. 1: Introduction. *Bmj, 353*, i2016.

Asdal, K. (2012). Contexts in Action—and the Future of the Past in Sts. *Science, Technology, & Human Values, 37*(4), 379-403.

Asdal, K., & Moser, I. (2012). Experiments in Context and Contexting. *Science, Technology, & Human Values, 37*(4), 291-306.

Bærøe, K. (2018). On Fundamental Premises for Addressing" Context" and" Contextual Factors" Influencing Value Decisions in Healthcare: Comment on" Contextual Factors Influencing Cost and Quality Decisions in Health and Care: A Structured Evidence Review and Narrative Synthesis". *International Journal of Health Policy and Management, 7*(10), 958.

Bærøe, K., & Baltussen, R. (2014). Legitimate Healthcare Limit Setting in a Real-World Setting: Integrating Accountability for Reasonableness and Multi-Criteria Decision Analysis. *Public Health Ethics, 7*(2), 98-111. doi: 10.1093/phe/phu006

Baltussen, R., Jansen, M. P., Mikkelsen, E., Tromp, N., Hontelez, J., Bijlmakers, L., & Van Der Wilt, G. J. (2016). Priority Setting for Universal Health Coverage: We Need Evidence-Informed Deliberative Processes, Not Just More Evidence on Cost-Effectiveness. *International Journal of Health Policy and Management, 5*(11), 615-618. doi: 10.15171/ijhpm.2016.83

Baltussen, R., & Niessen, L. (2006). Priority Setting of Health Interventions: The Need for Multi-Criteria Decision Analysis. *Cost effectiveness and resource allocation, 4*(1), 14. doi: 10.1186/1478-7547-4-14

Bazzoli, G. J., Clement, J. P., Lindrooth, R. C., Chen, H.-F., Aydede, S. K., Braun, B. I., & Loeb, J. M. (2007). Hospital Financial Condition and Operational Decisions Related to the Quality of Hospital Care. *Medical Care Research and Review, 64*(2), 148-168.

Blumer, H. (1954). What Is Wrong with Social Theory? *American sociological review, 19*(1), 3-10.

Booth, C. M., Dranitsaris, G., Gainford, M. C., Berry, S., Fralick, M., Fralick, J., . . . Clemons, M. (2007). External Influences and Priority-Setting for Anti-Cancer Agents: A Case Study of Media Coverage in Adjuvant Trastuzumab for Breast Cancer. *BMC cancer, 7*(1), 110.

Calnan, M. (2018). Decisions of Value: Going Backstage. *International Journal of Health Policy Management*, 1-7.

Castro, M. F., Guccio, C., Pignataro, G., & Rizzo, I. (2014). The Effects of Reimbursement Mechanisms on Medical Technology Diffusion in the Hospital Sector in the Italian Nhs. *Health Policy, 115*(2-3), 215-229.

Centre, B. H. C. K. (2017). Annual Report 2017. In B. H. C. K. Centre (Ed.). https://annualreport.kce.be/2017/en/annual-report-2017/.

Centre, B. H. C. K. (2019). Our Team.

Cerri, K. H., Knapp, M., & Fernandez, J.-L. (2014). Decision Making by Nice: Examining the Influences of Evidence, Process and Context. *Health Economics, Policy and Law, 9*(2), 119-141.

Cerri, K. H., Knapp, M., & Fernandez, J.-L. (2015). Untangling the Complexity of Funding Recommendations: A Comparative Analysis of Health Technology Assessment Outcomes in Four European Countries. *Pharmaceutical Medicine, 29*(6), 341-359.

Cleemput, I., Devriese, S., Kohn, L., & Westhovens, R. (2018). A Multi-Criteria Decision Approach for Ranking Unmet Needs in Healthcare. *Health Policy, 122*(8), 878-884.

Clement, F. M., Harris, A., Li, J. J., Yong, K., Lee, K. M., & Manns, B. J. (2009). Using Effectiveness and Cost-Effectiveness to Make Drug Coverage Decisions: A Comparison of Britain, Australia, and Canada. *Jama, 302*(13), 1437-1443.

College Voor Zorgverzekeringen. (2008a). Begeleiding Bij Stoppen Met Roken: Verzekerde Zorg?

College Voor Zorgverzekeringen. (2008b). Brief Aan De Minister Van Volksgezondheid, Welzijn En Sport, Vergoeding Benzodiazepinen

College Voor Zorgverzekeringen. (2009a). Pakketadvies 2009, Publicatienummer 274.

College Voor Zorgverzekeringen. (2009b). Stoppen-Met-Rokenprogramma: Te Verzekeren Zorg! Publicatienummer 276

College Voor Zorgverzekeringen. (2010). Rapport Hulpmiddelenzorg 2010, Publicatienummer 286.

Commissie-Dunning. (1991). Kiezen En Delen. *Rapport van de Commissie Keuzen in de Zorg*.

Couwenbergh, B., Van Der Meer, F., Weghaus-Reus, S., Schelleman, H., & Zwaap, J. (2013). Pakketbeheer in De Praktijk Deel 3 *Diemen: Zorginstituut Nederland* (pp. 1-106).

Creswell, J. W., & Poth, C. N. (2017). *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. USA: Sage publications.

Cromwell, I., Peacock, S. J., & Mitton, C. (2015). 'Real-World' Health Care Priority Setting Using Explicit Decision Criteria: A Systematic Review of the Literature. *BMC health services research, 15*(1), 164. doi: 10.1186/s12913-015-0814-3

Csanádi, M., Löblová, O., Ozierański, P., Harsányi, A., Kaló, Z., Mckee, M., & King, L. (2019). When Health Technology Assessment Is Confidential and Experts Have No Power: The Case of Hungary. *Health Economics, Policy and Law, 14*(2), 162-181.

Dakin, H., Devlin, N., Feng, Y., Rice, N., O'neill, P., & Parkin, D. (2015). The Influence of Cost-Effectiveness and Other Factors on Nice Decisions. *Health Economics, 24*(10), 1256-1271. doi: doi:10.1002/hec.3086

Daniels, N. (2000). Accountability for Reasonableness. *Establishing a fair process for priority setting is easier than agreeing on principles, 321*(7272), 1300-1301. doi: 10.1136/bmj.321.7272.1300

Daniels, N., Porteny, T., & Urrutia, J. (2016). Expanded Hta: Enhancing Fairness and Legitimacy. *International Journal of Health Policy and Management, 5*(1), 1-3. doi: 10.15171/ijhpm.2015.187

Daniels, N., & Sabin, J. (1997). Limits to Health Care: Fair Procedures, Democratic Deliberation, and the Legitimacy Problem for Insurers. *Philosophy & Public Affairs, 26*(4), 303-350. doi: 10.1111/j.1088-4963.1997.tb00082.x

Daniels, N., & Sabin, J. (1998). The Ethics of Accountability in Managed Care Reform. *Health Affairs, 17*(5), 50-64. doi: 10.1377/hlthaff.17.5.50

Daniels, N., & Sabin, J. (2008). Accountability for Reasonableness: An Update. *Bmj, 337*. doi: 10.1136/bmj.a1850

De Fine Licht, J. (2011). Do We Really Want to Know? The Potentially Negative Effect of Transparency in Decision Making on Perceived Legitimacy. *Scandinavian Political Studies, 34*(3), 183-201.

Deville, J., Guggenheim, M., & Hrdličková, Z. (2016). Practising Comparison: Logics Relations Collaborations.

Eddama, O., & Coast, J. (2008). A Systematic Review of the Use of Economic Evaluation in Local Decision-Making. *Health Policy, 86*(2-3), 129-141.

Franken, M., Le Polain, M., Cleemput, I., & Koopmanschap, M. (2012). Similarities and Differences between Five European Drug Reimbursement Systems. *International Journal of Technology Assessment in Health Care, 28*(4), 349-357. doi: 10.1017/S0266462312000530

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, 119*(2), 195-202. doi: https://doi.org/10.1016/j.healthpol.2014.10.007

G-Ba. (2016). Dossier Zur Nutzenbewertung Gemäß § 35a Sgb V, Nivolumab (Opdivo®).

G-Ba. (2017). *Verordnungsausschluss Von Arzneimitteln Zur Erhöhung Der Ebensqualität Gemäß § 34 Abs. 1 Satz 7 Sgb V (Lifestyle Arzneimittel), Anlage Ii Zum Abschnitt F Der Arzneimittel-Richtlinie Stand (Letzte Änderung in Kraft Getreten)*.

Gemeinsamer Bundesausschuss. The Federal Joint Committee: Who We Are and What We Do.

Giacomini, M., Hurley, J., & Stoddart, G. (2000). The Many Meanings of Deinsuring a Health Service: The Case of in Vitro Fertilization in Ontario. *Social Science & Medicine, 50*(10), 1485-1500. doi: https://doi.org/10.1016/S0277-9536(99)00394-9

Gkv-Spitzenverband. (2018). Fortschreibung Der Produktgruppe 10 „Gehhilfen“Des Hilfsmittelverzeichnisses Nach § 139 Sgb V. https://www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/hilfsmittel/fortschreibungen_aktuell/2018_2/20180827_Fortschreibung_der_Produktgruppe_10_Gehhilfen.pdf.

Grosios, K., Gahan, P. B., & Burbidge, J. (2010). Overview of Healthcare in the Uk. *EPMA Journal, 1*(4), 529-534.

Guyatt, G. H., Oxman, A. D., Kunz, R., Falck-Ytter, Y., Vist, G. E., Liberati, A., & Schünemann, H. J. (2008). Going from Evidence to Recommendations. *Bmj, 336*(7652), 1049-1051.

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). Grade: An Emerging Consensus on Rating Quality of Evidence and Strength of Recommendations. *Bmj, 336*(7650), 924-926.

Hall, W. (2017). Don't Discount Societal Value in Cost-Effectiveness: Comment on "Priority Setting for Universal Health Coverage: We Need Evidence-Informed Deliberative Processes, Not Just More Evidence on Cost-Effectiveness". *International Journal of Health Policy and Management, 6*(9), 543-545. doi: 10.15171/ijhpm.2017.03

Hasman, A., Mcintosh, E., & Hope, T. (2008). What Reasons Do Those with Practical Experience Use in Deciding on Priorities for Healthcare Resources? A Qualitative Study. *Journal of Medical Ethics, 34*(9), 658-663. doi: 10.1136/jme.2007.023366

Hoedemaekers, R., & Oortwijn, W. (2003). Problematic Notions in Dutch Health Care Package Decisions. *Health Care Analysis, 11*(4), 287-294. doi: 10.1023/B:HCAN.0000010057.43321.b2

Holstein, J. A., & Gubrium, J. F. (2016). *Narrative Practice and the Active Interview* (Vol. 67): Sage London.

Hughes, D., & Light, D. (2002). *Rationing: Constructed Realities and Professional Practices*. Oxford, UK: Wiley-Blackwell.

Ispor. (2009). Pharmaceutical Hta and Reimbursement Processes - Germany. In ISPOR (Ed.), *Global Health Care Systems Road Map*. https://www.ispor.org/HTARoadMaps/Germany.asp.

Jansen, M. P., Helderman, J.-K., Boer, B., & Baltussen, R. (2017). Fair Processes for Priority Setting: Putting Theory into Practice : Comment on "Expanded Hta: Enhancing Fairness and Legitimacy". *International Journal of Health Policy and Management, 6*(1), 43-47. doi: 10.15171/ijhpm.2016.85

Kankerplan, N. (2008-2010). Nationaal Kankerplan.

Kapiriri, L., Norheim, O. F., & Martin, D. K. (2009). Fairness and Accountability for Reasonableness. Do the Views of Priority Setting Decision Makers Differ across Health Systems and Levels of Decision Making? *Social Science & Medicine, 68*(4), 766-773. doi: https://doi.org/10.1016/j.socscimed.2008.11.011

Kleijnen, S., George, E., Goulden, S., D'andon, A., Vitré, P., Osińska, B., . . . Nagy, B. Z. (2012). Relative Effectiveness Assessment of Pharmaceuticals: Similarities and Differences in 29 Jurisdictions. *Value in Health, 15*(6), 954-960.

Kleinhout-Vliek, T., De Bont, A., & Boer, B. (2017). The Bare Necessities? A Realist Review of Necessity Argumentations Used in Health Care Coverage Decisions. *Health Policy, 121*(7), 731-744. doi: https://doi.org/10.1016/j.healthpol.2017.04.011

Kleinhout-Vliek, T., De Bont, A., & Boer, B. (2020). Necessity under Construction – Societal Weighing Rationality in the Appraisal of Health Care Technologies. *Journal of health economics, policy, and law*.

Krause, M. (2016). Comparative Research: Beyond Linear-Casual Explanation.

Lamont, M. (2012). Toward a Comparative Sociology of Valuation and Evaluation. *Annual Review of Sociology, 38*(1), 201-221. doi: 10.1146/annurev-soc-070308-120022

Lamont, M., & Thévenot, L. (2000). Introduction: Toward a Renewed Comparative Cultural Sociology. *Rethinking comparative cultural sociology: repertoires of evaluation in France and the United States*, 1-22.

Makady, A., Ten Ham, R., De Boer, A., Hillege, H., Klungel, O., & Goettsch, W. (2017). Policies for Use of Real-World Data in Health Technology Assessment (Hta): A Comparative Study of Six Hta Agencies. *Value in Health, 20*(4), 520-532.

Mann, A. (2015). Which Context Matters? Tasting in Everyday Life Practices and Social Science Theories. *Food, Culture & Society, 18*(3), 399-417.

Martin, D., Pater, J., & Singer, P. (2001a). Priority-Setting Decisions for New Cancer Drugs: A Qualitative Case Study. *The Lancet, 358*(9294), 1676-1681. doi: https://doi.org/10.1016/S0140-6736(01)06714-9

Martin, D. K., Pater, J. L., & Singer, P. A. (2001b). Priority-Setting Decisions for New Cancer Drugs: A Qualitative Case Study. *The Lancet, 358*(9294), 1676-1681.

Maynou, L., & Cairns, J. (2019). What Is Driving Hta Decision-Making? Evidence from Cancer Drug Reimbursement Decisions from 6 European Countries. *Health Policy, 123*(2), 130-139.

Miller, R., Williams, I., Allen, K., & Glasby, J. (2013). Evidence, Insight, or Intuition? Investment Decisions in the Commissioning of Prevention Services for Older People. *Journal of Care Services Management, 7*(4), 119-127.

Moes, F., Houwaart, E., Delnoij, D., & Horstman, K. (2017). Contested Evidence: A Dutch Reimbursement Decision Taken to Court. *Health Economics, Policy and Law, 12*(3), 325-344.

Moreira, T. (2011). Health Care Rationing in an Age of Uncertainty: A Conceptual Model. *Social Science & Medicine, 72*(8), 1333-1341.

Nguyen-Kim, L., Or, Z., Paris, V., & Sermet, C. (2005). The Politics of Drug Reimbursement in England, France and Germany. *Issues in health economics, 99*.

Nhs. (2018). Walking Aids, Wheelchairs and Mobility Scooters. *2019*.

Nice. Generalised Anxiety Disorder in Adults, the Nice Guideline on Management in Primary, Secondary and Community Care.

Nice. (2011). Generalised Anxiety Disorder and Panic Disorder in Adults: Management. Clinical Guideline [Cg113].

Nice. (2013). Guide to the Methods of Technology Appraisal.

Nice. (2016). Technology Appraisal Processes - Cdf. https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICEtechnology-appraisals/process-and-methods-guide-addendum.pdf.

Nice. (2017a). Nice Technology Appraisal Guidance: Nivolumab for Previously Treated Squamous Non-Small-Cell Lung Cancer (Ta483).

Nice. (2017b). Nivolumab Now Available for Lung Cancer after Company Offers Nice New Cdf Deal. *2019*(7 February 2019). https://www.nice.org.uk/news/article/nivolumab-now-available-for-lung-cancer-after-company-offers-nice-new-cdf-deal

Nice. (2018). Stop Smoking Interventions and Services, Nice Guideline (Ng92).

Nicod, E., & Kanavos, P. (2012). Commonalities and Differences in Hta Outcomes: A Comparative Analysis of Five Countries and Implications for Coverage Decisions. *Health Policy, 108*(2-3), 167-177.

Overheid, V. (2019). Rookstopbegeleiding. Retrieved 8 March 2019, from https://www.vlaanderen.be/nl/gezin-welzijn-en-gezondheid/gezond-leven/rookstopbegeleiding

Patera, N., & Wild, C. (2014). Assessment – Appraisal – Decision. Lbi-Hta Decision Support Document Nr.: 72. Wien: Ludwig Boltzmann Institut für Health Technology Assessment.

Peacock, S., & Bentley, C. (2018). Understanding Contextual Factors in Cost, Quality and Priority Setting Decisions in Health: Comment on" Contextual Factors Influencing Cost and Quality Decisions in Health and Care: A Structured Evidence Review and Narrative Synthesis". *International Journal of Health Policy and Management, 7*(12), 1145.

R.I.Z.I.V. (2016). Evaluatierapport Dag 90 Dossier: 1542 Opdivo - Tweede Aanvraag - Version 06 07 2016 In D. v. d. G. Verzorging (Ed.).

Ragin, C. C. (2004). Turning the Tables: How Case-Oriented Research Challenges. *Rethinking social inquiry: Diverse tools, shared standards, 123*.

Riziv-Inami. (2003). Nota Hoofdwerkgroep Nr. 2003/6.4, Hoofdgroep 1.4 Loophulpen Volwassenen. In D. g. v.-S. d. s. d. santé (Ed.).

Rooshenas, L., Owen-Smith, A., Hollingworth, W., Badrinath, P., Beynon, C., & Donovan, J. L. (2015). "I Won't Call It Rationing…": An Eth-

nographic Study of Healthcare Disinvestment in Theory and Practice. *Social Science & Medicine, 128*(Supplement C), 273-281. doi: https://doi.org/10.1016/j.socscimed.2015.01.020

Rosenberg-Yunger, Z. R., Daar, A. S., Thorsteinsdóttir, H., & Martin, D. K. (2011). Priority Setting for Orphan Drugs: An International Comparison. *Health Policy, 100*(1), 25-34.

Salas-Vega, S., Bertling, A., & Mossialos, E. (2016). A Comparative Study of Drug Listing Recommendations and the Decision-Making Process in Australia, the Netherlands, Sweden, and the Uk. *Health Policy, 120*(10), 1104-1114.

Schippers, E. (2015). 29 477 Geneesmiddelenbeleid. In W. e. S. Ministerie van Volksgezondheid (Ed.). https://zoek.officielebekendmakingen.nl/kst-29477-343-n1.html.

Singer, P. A., Martin, D., Giacomini Bhatia, M. V., & Purdy, L. (2000). Priority Setting for New Technologies in Medicine: Qualitative Case Study. *Bmj, 321*(7272), 1316-1318. doi: 10.1136/bmj.321.7272.1316

Stolk, E., Brouwer, W., & Busschbach, J. (2002). Rationalising Rationing: Economic and Other Considerations in the Debate About Funding of Viagra. *Health Policy, 59*(1), 53-63.

Stolk, E., Goes, E., Kok, E., & Busschbach, J. (2001). Uitwerking Criteria Noodzakelijkheid, Eigen Rekening En Verantwoording En Lifestyle. *College voor Zorgverzekeringen Breedte geneesmiddelenpakket. Amstelveen, The Netherlands: College voor Zorgverzekeringen, 1*, 54.

Van Den Bruel, A., Cleemput, I., Van Linden, A., Schoefs, D., Ramaekers, D., & Bonneux, L. (2004). Effectiviteit En Kosten-Effectiviteit Van Behandelingen Voor Rookstop. In K. v. d. G. (KCE) (Ed.).

Vuorenkoski, L., Toiviainen, H., & Hemminki, E. (2008). Decision-Making in Priority Setting for Medicines—a Review of Empirical Studies. *Health Policy, 86*(1), 1-9.

Williams, I., Brown, H., & Healy, P. (2018). Contextual Factors Influencing Cost and Quality Decisions in Health and Care: A Structured Evidence Review and Narrative Synthesis. *International Journal of Health Policy and Management, 7*(8), 683.

Williams, I., Brown, H., & Healy, P. (2019). Influencing Decisions of Value in Health: A Response to Recent Commentaries. *International Journal of Health Policy and Management, 8*(3), 187.

Williams, I., Mciver, S., Moore, D., & Bryan, S. (2008). The Use of Economic Evaluations in Nhs Decision-Making: A Review and Empirical Investigation. *HEALTH TECHNOLOGY ASSESSMENT-SOUTHAMPTON-, 12*(7).

Wirtz, V., Cribb, A., & Barber, N. (2005). Reimbursement Decisions in Health Policy—Extending Our Understanding of the Elements of Decision-Making. *Health Policy, 73*(3), 330-338.

Zorginstituut Nederland. (2013). Pakketbeheer in De Praktijk Deel 3 (pp. 1-106). Diemen.

Zorginstituut Nederland. (2015). Pakketadvies Nivolumab (Opdivo) Incl. Brief Aan De Minister Voor Volkgezondheid, Welzijn En Sport.

Zorginstituut Nederland. (2016). Reglement Adviescommissie Pakket. In Z. Nederland (Ed.). https://www.zorginstituutnederland.nl/over-ons/publicaties/besluit/2016/12/05/reglement-adviescommissie-pakket-zorginstituut-nederland (accessed 11 February 2020).

Zorginstituut Nederland. (2017). Pakketadvies in De Praktijk: Wikken En Wegen Voor Een Rechtvaardig Pakket. Diemen.

3 | Expertise in the appraisal phase

Modified version of paper accepted to Journal of Health Economics, Policy and Law as: *Kleinhout-Vliek, T., de Bont, A., & Boer, B. Necessity under Construction – societal weighing rationality in the appraisal of health care technologies*

INTRODUCTION

Whether a health care technology is made available to patients, for example through a (national) formulary, benefits package, or insurance scheme, is decided by means of health care coverage decisions. Making these decisions well is notoriously complex as it leans on a wide variety of heterogeneous considerations that are brought together in the decision-making process (Cerri et al., 2014; Dakin et al., 2015; Hughes & Doheny, 2011; Russell & Greenhalgh, 2014; Vuorenkoski et al., 2008). The most-studied part of this process is the assessment, which examines the available scientific evidence such as the (cost-) effectiveness of the technology, often employing Health Technology Assessment (HTA) methods (eg. Franken et al., 2012; Kleijnen et al., 2012; Le Polain et al., 2010; Makady et al., 2017; Salas-Vega et al., 2016). Notably, such scientific knowledge bases have been shown essential but not sufficient for making good coverage decisions (Calnan et al., 2017; Moes et al., 2016; Samenleving, 2017; Shirazi et al., 2017). Therefore, many coverage decisions feature a second part, sometimes (but not always) a distinct step in time and space: the appraisal. An appraisal entails the formulation of a (recommended) coverage decision based on a contextualisation of the given evidence (Jansen et al., 2017; Kleinhout-Vliek et al., 2020; Oliver et al., 2004; Patera & Wild, 2014; Walley, 2007). Previously, contextualisation has broadly been defined as taking into account a variety of values and considerations (Patera & Wild, 2014). Some, such as the English National Institute for Health and Care Excellence (NICE), have specified contextualisation to mean establishing "what is good for society" or, in other words, providing a societal weighing of the evidence (Culyer & Rawlins, 2004; NICE, 2008; Shah et al., 2013).

Little is known about how societal weighing works and how exactly appraisal committees may achieve this. Recent work on deliberative coverage decisions more generally does show that expertise in these decisions may comprise three distinct elements. First, the expertise is substantive; committee members are able to understand and work with different types of knowledge. These types include not only the scientific input deriving from the assessment but often also knowledge provided by patients ('experts-by-experience') as well as 'local' knowledge regarding the institutional setting where the decision-making process takes place. Second, committee members respond appropriately to these different types of input and combine them into one decision, which involves being 'rational' and being 'human'. They also deal with any tensions that may arise between knowledge types (Calnan et al., 2017; Gkeredakis et al., 2011; Hughes & Doheny, 2011; Jenkings & Barber, 2004; Moes et al., 2016; Moreira, 2011; Russell & Greenhalgh, 2014). Third, committee members are experts at adhering to not only substantive requirements which are generally in the form of formalised decision criteria but also processual requirements (Russell, 2017), often those laid out by the widely-used Accountability for Reasonableness framework (Daniels, 2000; Daniels & Sabin, 1997, 1998, 2008; Hasman & Holm, 2005). These three elements of expertise would appear, however, not specific to appraisals so the question remains: what does societal weighing expertise look like?

In this study, we will fill the gap concerning societal weighing expertise by examining the appraisal phase of Dutch health care coverage decisions, which also, like NICE's decisions described

above, explicitly aim for a "societal weighing" of the evidence base provided in the assessment (Zorginstituut Nederland, 2016, 2017). Furthermore, Dutch decision-making practice provides an excellent case because the appraisal is set up as a separate meeting, a distinct moment in time and space, making it easier to distinguish from the assessment phase (Commissie-Dunning, 1991; Stolk et al., 2002; Stolk et al., 2001). In the Netherlands, the National Health Care Institute ('the Institute', in Dutch: *Zorginstituut Nederland*), an arm's length body, is responsible for advising the Ministry of Health on the contents of the basic benefits basket. This basket outlines the bare-minimum health care insurance package that is obligatory for all Dutch citizens. The Dutch coverage decision-making process usually starts with a scoping session, in which stakeholders are invited to contribute relevant considerations. This is followed by a meeting of the Scientific Advisory Committee ('the assessment committee', in Dutch: *Wetenschappelijke Adviesraad*), in which the scientific knowledge base is established based on HTA methodology (Franken et al., 2012; Kleijnen et al., 2012; Makady et al., 2017; Salas-Vega et al., 2016). The Package Advisory Committee ('the appraisal committee', in Dutch: *Adviescommissie Pakket*), convened once-monthly at the Institute, is subsequently responsible for the societal weighing of the evidence and the formulation of coverage advice to the Minister (Couwenbergh et al., 2013; Zorginstituut Nederland, 2016, 2017). The Institute uses four formalised criteria in this advice, namely effectiveness, cost-effectiveness, feasibility (including budget impact considerations), and necessity. 'Necessity' is, as a consequence of being a formalised criterion, present in the documents providing input for the appraisal as they outline the available information per criterion. It has, however, also been considered to be established especially in appraisal (Couwenbergh et al., 2013; Kleinhout-Vliek et al., 2017; Kleinhout-Vliek et al., 2020; Mastenbroek et al., 2006; Zorginstituut Nederland, 2017). Therefore, this chapter will explore how the appraisal committee *constructs necessity* as this may aid in uncovering elements that are specific to societal weighing expertise (Hoedemaekers & Oortwijn, 2003; Poley, 2002; Stolk et al., 2002).

Aim

This chapter describes the phases of constructing necessity by examining the contents of and the proceedings at Dutch appraisal meetings through observations, transcriptions, and subsequent analysis of audio recordings of four meetings. These meetings varied greatly in terms of type of health care technology appraised. This is supplemented with interviews with appraisal committee members and Institute employees (n=13). We answer the question: how does the Dutch appraisal committee construct necessity?

METHODOLOGY

Approach

In order to explore the Dutch appraisal, we analyse four cases of specific coverage decisions. Argumentations used to construct necessity that could potentially be employed by the appraisal commit-

tee include but are not limited to, the morbidity and need experienced by the patient, availability of alternative treatments, the financial cost per individual patient, and compassion with vulnerable groups such as children or small numbers of patients generally. These necessity argumentations are part of a list of twenty argumentation types derived from a realist review of argumentations used to establish the necessity of coverage of health care technologies worldwide (Kleinhout-Vliek et al., 2017). For an overview of these twenty inductively-formulated argumentation types, see Table 1. The Netherlands is a particularly fruitful setting for studying necessity as the Dutch use two of these twenty necessity argumentations, namely individual severity of illness ('Morbidity/Severity') (Franken et al., 2015; Reckers-Droog et al., 2018, 2019) and the cost that the individual patient will incur ('Individual Cost'), as their explicit, *a priori* formulated, necessity *criterion*. As such, these ought to be present in every decision document inputting into the appraisal (Couwenbergh et al., 2013; Hoedemaekers & Oortwijn, 2003; Niëns, 2014; Stolk et al., 2001).

Necessity argumentations are of interest for two reasons. First, these argumentations are employed not only by professional decision makers but also by other parties, such as patients, who may also be present at the appraisal meeting (further information below). Second, necessity argumentations are variable in usage as their perceived validity and allotted weight differs per decision, making their patterns especially vivacious (Kleinhout-Vliek et al., 2017; Kleinhout-Vliek et al., 2020).

To study necessity argumentations in appraisal we have chosen a case approach (Creswell & Poth, 2017; Ragin, 2004) because researching a similar process in a variety of situations is considered helpful for gaining insight into commonalities across situations, especially when it concerns context-dependent social processes (Dussauge et al., 2015; Lamont, 2012). In this, we hold that the context is case-specific and thus differs per decision situation (Asdal, 2012; Asdal & Moser, 2012). The cases chosen concern four health care technologies, namely eculizumab (Soliris®), front teeth replacement, maternity care, and paracetamol-vitamin D tablets. These cases are relatively run-of-the-mill; three were discussed in the Dutch media but only sparingly (Berkhout, 2017; Unknown, 2015, 2016; Van der Aa, 2016). They share two characteristics that resulted in their selection. First, the cases vary widely, i.a. in terms of type of technology, price, and number of patients affected, with necessity argumentations playing a pivotal role in all four, as suggested in informal conversations by contacts at the Institute (front teeth replacement therapy, maternity care, and paracetamol-vitamin D tablets) or the literature review (eculizumab (Soliris®)) (Kleinhout-Vliek et al., 2017). Second, the appraisal meetings took place between 2015 and 2017, meaning that the first author could either be present at the appraisal meeting, listen to the audio files, or do both (see Table 2).

Methods

Our dataset comprises four cases and the data on these cases were gathered through observations, transcription of audio files, documents, and interviews. The Institute consented to the first author accessing the setting of appraisal committee through an explicit invitation to do so by the secretary of the committee. This invitation included the 'closed' pre-meeting for the eculizumab and the paracetamol-vitamin D cases and additional observations at the scoping session for the latter. The

Argumentation type	Description
Definition of Illness	Whether the ailment is considered an illness for which treatment is necessary
Dignity	Whether (lack of) coverage is considered to affect the dignity of the patient to such an extent that it needs to be amended
Equity/Fairness/Justice	Whether coverage would be necessary to counter injustice/inequity/lack of fairness in (access to) treatment
Human Right	Whether (lack of) coverage is considered to affect the human rights of the patient to such an extent that it needs to be amended
Individual Cost	Whether lack of coverage would stop patients from buying necessary care themselves due to prohibitive cost
Individual Responsibility	Whether the individual is considered responsible for paying for this treatment
Medical Necessity	Whether or not a treatment is considered to be "medically necessary" or a "medical necessity"
Morbidity/Severity	Whether the physical and/or psychosocial morbidity associated with a certain ailment constitutes such a need that coverage is considered necessary
Moral Hazard	Whether there is considered to be a possibility of over-usage (i.e., unnecessary increase in demand, when people use more than they need as a result of coverage)
Need	The extent to which the patient is considered to be in need for which treatment is necessary
(No) Alternative	Whether or not viable alternatives are considered to be present which would make coverage more or less necessary
Patient-Diagnosis	Whether an illness is self-reported rather than diagnosed by a doctor
Range of Normality	Whether the experience of the patient is considered normal or abnormal to such an extent that coverage is deemed necessary
Rule of Rescue	Whether the identifiability of individuals close to death is considered to heighten the necessity of coverage
Similar Treatments	Whether similar treatments are covered or not (meaning that this type of treatment is considered necessary)
Small Number of Patients	Whether the small size of the patient population is considered to heighten the necessity of coverage (due to, amongst others, the inequality in terms of research expenditure or difficulties in obtaining high-quality data)
Societal Impact	Whether coverage is considered necessary to allay the impact this disease has on people beyond the patient
Societal Functioning	Whether coverage would aid a person's necessary functioning in society
Societal Responsibility	Whether society is considered responsible for paying for this treatment
Vulnerability/Compassion	Whether a compassionate response to vulnerable groups, e.g. children, in the form of coverage is considered to be a necessity

Table 1- Overview of the twenty argumentation types that fall under the necessity criterion and their respective descriptions (Kleinhout-Vliek et al., 2017)

secretary also provided audio-recordings s/he herself used to write minutes to be analysed for this study. The first author was present at three out of four appraisal committee meetings (for the cases eculizumab, front teeth replacement therapy, and paracetamol-vitamin D tablets), where she observed and took field notes. These notes were supplemented by audio files of the same appraisal committee meetings and one additional case (maternity care). These audio files were transcribed verbatim. Moreover, we analysed nine documents pertaining to these cases. For each case, this entailed the document that was provided to the appraisal committee (the 'discussion document')

and the final 'appraisal report' (see Table 2). For the documents, only the main body of text was analysed (i.e., excluding appendices).

For triangulation purposes, the first author interviewed seven policy advisers ('Institute employees') who worked on the cases as well as six appraisal committee members present at the meetings, some of whom were interviewed multiple times, to a total of thirteen people over twelve interviews (for a precise overview of who was interviewed when, please see Table 2). All approached interviewees consented to being interviewed, except one Institute employee, who declined due to a full

Case	Observations	Document analysis	Interview
Eculizumab (Soliris®)	file 1 (October 2016)	3.1 discussion document 3.2 appraisal report 3.3 patient contribution	-
Front teeth replacement	file 2 (February 2015)	2.1 discussion document 2.2 appraisal report	Institute employee 3 (April 2015) Institute employee 4 (May 2015)
Maternity care	file 3 (January 2015) *Audio file only*	1.1 discussion document 1.2 appraisal report	Institute employees 2 & 5 (October 2016)
Paracetamol-vitamin D tablets	file 4 (October 2016)	4.1 discussion document 4.2 appraisal report	Institute employees 1 & 7 (October 2016) Committee member 5 (October 2016)

Table 2 – Overview of data collected pertaining to the case studies

schedule. Three interviews were group interviews (type: field-formal, meaning that the questions were of a semi-structured nature and the interviewer took on a somewhat directive role) (Frey & Fontana, 1991). Institute employees 2 & 5 and 1 & 7 were interviewed in pairs at the request of the employees themselves, as they considered their answers would supplement one another (in the former case, the interviewer did not know two people would be present until the moment of the interview). The group interview with Committee members 1, 4, 5 & 6 was done out of convenience, as committee members are often only present at the Institute once a month and this presented a good opportunity. Again, the interviewees considered their answers to complement one another. Oral informed consent was given for use of interview data, written informed consent was given for publication, and a formal waiver for ethical approval was obtained [reference number to be added]. The interviews were conducted by the first author using a topic list with non-structured, open-ended questions, and the interviews were audio-recorded and subsequently transcribed verbatim. The topic list included for Institute employees: how this technology arrived at the Institute agenda, the scoping session, how they retrieved any additional information, the appraisal meeting, how they arrived at the text in the different documents, and more general topics pertaining to the Institute. This was done to gain deeper insight into the working processes at the Institute, especially concerning different forms of input for the appraisal meeting. For committee members, the topic list concerned the appraisal committee's functioning generally, how different types of information

are usually dealt with, and specific experiences they could recall. These questions were formulated in order to gain insight into the cases but also to obtain reflections on tentatively formulated phases of necessity construction. Lastly, all interviewees were probed about necessity argumentations.

This dataset was analysed as follows. The list of twenty necessity argumentations (Kleinhout-Vliek et al., 2017) was used as sensitising concepts to guide the first step of the detailed content analysis of the documents and transcribed committee meetings, in which the necessity argumentations were used as a list of codes (Table 1). The explicit mentioning, 'black on white' or 'out loud', of necessity argumentations was tracked across the documents and audio files through coding specific utterances as one or more of the argumentation types, using Microsoft Excel to put utterances that had received the same code together in the same row. The first tentative patterns were subsequently elucidated based on this tracked argumentation use. For the formulation of the phases, we chose a chronological and person-dependent (first this person spoke, then that person contributed) rather than substantive (these argumentations were used more often than those) patterning. These patterns formed the basis for the different phases, formulated and refined in further extensive discussions within the authorship team, then supplied and solidified by information retrieved from the interviews. These interviews specifically clarified the dynamics around the Institute employees' and the patient and patient representatives' contributions. Additional observations served to see whether data saturation was achieved (Table 3). A member check with the appraisal committee and several personal communications (committee meetings of December 2016 and February 2017, member check through a presentation to the committee on 14 April 2018, personal communications with committee members 5, 6 and 7) to see whether the interpretation made by the authors stayed close enough to the interpretation of those observed (Sayer, 2011) (see also Table 3). Especially the member check and the personal communications have positively impacted the reliability of the study in this regard; the personal communications followed the member check to clarify the interpretation of our data in a number of places, these are indicated in the text through reference to '(interviewee X, personal communications)'.

RESULTS

This section first describes the working procedure of the Institute and the appraisal committee, succeeded by an introduction to the case studies and descriptions of the way in which necessity is constructed per phase of the appraisal meeting.

Working procedure

The Institute's working procedure for formulating an advised decision follows the general assessment-appraisal pattern (Patera & Wild, 2014). Agenda setting varies and may happen through pharmaceuticals gaining market access or by another party, such as the Minister of Health or a professional organisation. Once it is placed on the agenda, one or two Institute employees take

Observations and audio file	Interview
December 2016	Institute employee 6
Observations only	(March 2015)
February 2017	Committee member 5
Observations only	(March 2015)
	Committee member 3
	(August 2015)
	Committee member 2
	(September 2015)
	Committee members 1, 4, 5 & 6
	(February 2017)
	Committee member 1
	(February 2017)
	Committee member 6
	(October 2017)

Table 3 – Overview of additional data collected

responsibility for this dossier; in this study, these were different people for each case. After a scoping session with interested parties, the scientific evidence reports are written by other Institute employees with expertise in therapeutic value, cost-effectiveness, and budget impact. These are bundled and combined with a short explainer by the one or two Institute employees who hold final responsibility for this dossier to benefit the assessment phase. The assessment phase takes place at the Scientific Advisory Committee (in Dutch: *Wetenschappelijke Adviesraad*, WAR), based on which an assessment report is composed by the secretary of this committee. This report provides a summative conclusion on the valuation, the size, and the probability of the effect of the medicine. This assessment report is sent to the stakeholders for consultation and consequently combined with input from the scoping session as well as a fresh explainer into a 'discussion document' by those responsible for this dossier. This is aided by the secretary of the assessment committee and approved by the secretary and chair of the appraisal committee, to benefit the appraisal phase (Committee member 5, personal communications) (Zorginstituut Nederland, 2017).

The appraisal subsequently takes place at the meeting of the Package Advisory Committee (in this text: the appraisal committee). The committee is comparatively small (Patera & Wild, 2014): eight to ten external experts (e.g. in medical ethics, pharmaco-economics, or medicine), who are not employed by the Institute, comprise the committee. Like NICE's Social Value Judgements, the Dutch appraisal explicitly aims for a societal weighing of the provided scientific evidence (Couwenbergh et al., 2013; Zorginstituut Nederland, 2016, 2017).

The committee members read the assessment report in advance of the meetings. All meetings, which are in principle open to the public, were preceded by a 'closed' meeting, in which patients and their representatives were absent but the Institute employee(s) responsible for the dossier were present, and the files are already discussed (observations February, November 2015, October, December 2016, and February 2017, see also Zorginstituut Nederland, 2016).

Cases

We studied the appraisal deliberations for four significantly varying cases.

Eculizumab (Soliris®) is an orphan drug currently licensed for Atypical Hemolytic Uremic Syndrome (aHUS) and Paroxysmal Nocturnal Hemoglobinuria (PNH). After four years of provisional coverage, a temporary coverage arrangement, the final advised decision to the Minister was to be drafted by the Institute in 2016. The discussion document (document 3.1) states that there is debate on whether the incremental cost-effectiveness ratio (ICER) for treatment of aHUS with eculizumab approximates the reference value at €80,000 per QALY for severe diseases. The calculated cost-effectiveness ratio was thus considered "highly unfavourable", but clinicians and patient organisations had initiated independent research on shortening the treatment period through improved start-stop criteria, which was expected to result in a more favourable ICER. A grant was already obtained for this research (though not for the medicine itself). After the formal presentation by the Institute employee responsible came the contributions of one patient and two patient representatives (in this case, the mother of a patient and a clinician). Especially the mother's emotional contribution was followed by an extended silence on the part of the committee, and many committee members vocalised their appreciation of these contributions. For the deliberations, the research on the new protocol was the primary focal point. The committee thought investing in this a worthy cause; the final advised decision stressed that the committee considered the initiative so commendable that it needed to remain possible to reimburse eculizumab from public funds within the research protocol (Document 3.2). Relief on the part of the patients was palpable; the chair suggested the committee take a break, and the patient (representative)s were congratulating each other, also a few committee members offered their congratulations (observations/audio file 1).

Front teeth replacement was discussed in the appraisal committee in February 2015 after the College of Dentists (in Dutch: *College van Adviserend Tandartsen*) placed it on the agenda through the contacts that Institute employees 3 and 4 had with them. The reason provided was that current legislation was perceived as a perverse stimulus with the situation being as follows. All dental care is covered by the Dutch basic benefit package until the insured's 18th birthday but not afterwards. This means that when young people lose their front teeth or were born without them, they may prefer to have them replaced before their 18th birthday (as the costs are approximately €3,500 for front teeth implants), whereas it is often better to do so later as the oral cavity is not fully grown until the age of 22. In the appraisal committee meeting where the coverage of front teeth implants was discussed, the topic was not considered of major importance or interest, even a little laughable, for its small budget impact (audio file 2 and interview with Institute employee 5). Institute employees 3 and 4 were especially aware of its political history; one regarded it a mistake that could have been prevented that current legislation did not specify the extended coverage until the age of 22. The discussions in the appraisal committee seemed relatively straightforward, with everybody largely leaning in the direction of extending coverage, until one Committee member apparently "wanted to stimulate the discussion by deliberately going against the tide" (interview Institute employee 3). This resulted in a longer discussion, with the final decision apparently taken for pragmatic reasons

(namely the time it would take to change the legislation; coming back to it in the appraisal meeting next month would mean another year's extension) (observations/audio file 2). The final advised decision, then, was that as long as the claim was made before the 18th birthday, coverage would be continued until the age of 22 (Document 2.2).

The maternity care case was addressed by the appraisal committee in January 2015 in response to an appeal to the general public by the Minister of Health to send in suggestions for parts of the basic benefit package that could be removed. The primary argumentation for suggesting that maternity care could be removed from the package was that pregnancy is a choice and foreseeable (that is, unnecessary to be insured for, as the birth is something you know will happen). In line with this, mere days before this appraisal committee meeting took place, a newspaper had commented on how strange it was that "beschuit met muisjes smeren" (the preparation of a traditional Dutch treat for friends and family paying maternity visits) was often done by maternity care workers and therefore part of the basic benefit package. The Institute employees, however, concurrently received "signals" from the Ministry that it should not be removed from the benefits basket (Institute employees 2 and 5, interview). Several versions of the advice document had to be discussed at the appraisal committee, which led to much frustration at coffee machine afterwards. The final document, however, was the first where the criteria were used "really well" (interviews with Institute employees 6 and 5). The deliberations in the appraisal meeting focused on two elements. First, "the domain question" (audio file 3): to what extent is maternity care a type of curative care? The second element concerned the idea of whether the need for maternity care is "foreseeable", meaning that if you know you are going to need something someday, you should not be *insured* for it (as you can save up, you know it is coming). The final decision was indeed to continue maternity care coverage (Document 1.2).

The paracetamol-vitamin D tablets file came onto the agenda of the Institute due to some national policy changes, by which a fairly large number of tablets was left behind on a list of covered medications (in Dutch: *Geneesmiddelenvergoedingssysteem*). In a letter to the Institute, the Minister of Health explicitly stressed the need to pay attention to the necessity of coverage of these tablets, which included 1000 mg paracetamol, vitamin D, and calcium tablets. The Institute employees responsible acknowledged that they, at first, called it an "outflow advice" but were "not allowed" to call it that (interview with Institute employees 1 and 7). It was discussed at the appraisal committee in October and November 2016. A number of pharmacists had inputted, as patient representatives, to the scoping session; the secretary of the committee, brought in the argumentations supplied by them. As a consequence, much of the deliberations focused on what constitutes individual affordability for specific vulnerable groups. The second element that was primary in the discussions was the comparator medicines; if we compare it to other medicines available at the drug store without a prescription, it should not be reimbursed. As a corollary, the question was whether these medicines constituted "self-care" medicines or not. The final advised decision was to not cover paracetamol-vitamin D tablets except for certain medical indications. The rationale was that reimbursing would actually be more expensive than not, because of the prescription rule. This rule is an extra charge

levied when a reimbursed medicine is bought at the pharmacy on prescription, which was circumvented by the negative advised decision (Document 4.2).

General setting

> The meetings of the appraisal committee take place in a sizeable meeting room that is relatively light, even though the blinds are drawn. There are two entrances: you can enter the room from 'within' the building (behind the security gates) but also from the 'outside', provided your name is on the list, which is checked at the reception. Ten people are seated around tables set up in a large square. There are thermoses with coffee and tea and plates of biscuits. An 'audience' of eight more people, me included, sit on the rows of chairs set up on one side of the room, where we can see the committee and the presentation screen well. I seem unable to shake the feeling we are watching a staged performance. It is clearly one of these occasions where you feel careful about making noise: I open my water bottle as quietly as possible. Given this fairly formal setting, I am struck every time by how warmly the committee members greet one another when they come in, how at-home they seem (one even brought her dog!), and even more by the apparent light-heartedness of it all, the sheer good humour that characterises the proceedings. (Condensed field notes 2, 3, 4)

In this setting, the deliberations on the four cases followed approximately the same order. We have separated this order out into four phases, namely 1) the contribution of the Institute employee(s); 2) the contributions of the patient(s) and/or their representative(s); 3) the actual deliberations of the committee; and 4) the formulation of the decision. This separation into four phases allows us to show how necessity is constructed in each phase.

Phase 1: Institute employee(s)

The contribution of the Institute employee(s) is the first of four phases we distinguish in the Dutch appraisal meeting, and we will show the impact of these contributions on the committee deliberations that followed.

The first phase of the 'open' meeting that followed, after the meeting was formally opened by the chair, comprised the contributions by the Institute employee(s). The individual severity of illness (code: Morbidity/Severity) and costs per the individual patient (code: Individual Cost) are formulated as official, explicit elements of the formalised necessity criterion. These were contributed by the Institute employee in the form of necessity argumentations, often by means of a formal presentation (observations 1, 2 and 4; audio files 1-4). Generally speaking, the appraisal committee took the individual severity of illness and costs per individual patient as 'given' in their deliberation; they did not explicitly mention them in their deliberations (observations 1, 2, 4/audio files 1-4). This pattern was broken only in the paracetamol-vitamin D case, in which the low severity of illness was indeed actively questioned by a committee member but not discussed any further.

The applicability of the individual cost argumentation was also questioned explicitly on behalf of the patient representatives by a Committee member, which did prompt a response that served to re-establish its applicability in general. Neither argumentations were 'weighed' compared to other argumentations.

As a committee member explained, the reason for this lack of use of the severity of illness and costs for the individual by the committee is that it is not their mandate to weigh these explicitly (Committee member 4, personal communication). This does not mean the Institute employee's input is ineffective. One committee member commented on the introduction by Institute employees, that:

> CM5 [T]he discussion, at some point, heads into a different direction.
> CM6 Yes, that's right. (Group interview with Committee members 1, 4, 5 and 6)

They acknowledged the impact on the deliberation: the Institute employee's argumentations deriving from formalised criteria are considered authoritative and steering the direction the committee's discussion takes.

In some cases, it could be argued that the appraisal committee did give a different formulation to these two explicit necessity criteria, that is, operationalising them differently (resulting in a different argumentation type code). This would involve not calling the individual severity of illness as between "0.71 and 1.00" (code: Morbidity/Severity, document 3.1) but instead, highlighting the daily consequences for the patient (code: Societal Functioning, observations/audio file 1). A similar reframing was observed in the front teeth case:

> [Just consider] what that means for someone, right, missing front teeth, in daily communication. (Observations/audio file 2, code: Societal Functioning)

This clearly shows how the committee may sometimes employ a different operationalisation, resulting in a different argumentation type code, of the criteria brought in by the Institute employee.

In sum, necessity is thus constructed in this first phase not by weighing formalised criteria explicitly but by rephrasing them and/or allowing them to steer the appraisal process implicitly, which committee members consider to positively impact the deliberations.

Phase 2: patient (representative)(s)

For the second phase of the appraisal, which consists of the contribution(s) of the patient (representative)(s), we will also describe how they affect the deliberations of the committee.

In the second phase, patients and/or patient representatives such as medical doctors, gave a short statement (observations/audio file 1). These people differ per case and are not always present. The observations/audio file data show that necessity argumentations that a patient (representative)

contributed were hardly ever mentioned by the appraisal committee. In the eculizumab case, the patient representative mentioned that the decision was:

> [A] story (…) that concerns (…) justice. (Document 3.3 and observations/audio file 1, code: Equity/Fairness/Justice)

> Not only is [coverage] the best option for the doctors and us; it is also for society as a whole. (Document 3.3 and observations/audio file 1, code: Societal Impact)

Neither of these argumentation types were mentioned as such by the appraisal committee (observations/audio file 1). The patient representative also brought in two argumentation types that were repeated once but not discussed further. First:

> As of [29 October 2014], the life of our little daughter Rosa, just 1 year old, forever lost its ease and was never again taken for granted. (Document 3.3 and observations/audio file 1, code: Vulnerability/Compassion)

This clear call for compassion was repeated by Committee member 7 (observations/audio file 1) but not discussed by the committee. The same goes for the functioning of the family generally, which the patient representative described as a "roller coaster" due to the high uncertainty of coverage. This element of uncertainty was mentioned in turn by the same committee member but also not discussed further (observations/audio file 1).

Though patients were only physically present in this dataset in the eculizumab case, their contributions were a primary topic of discussion during the interviews. Several committee members commented on what they experience when faced with patients and their representatives during the appraisal, and 'Distance' was the most important aspect identified:

> We need to keep the distance [from the patients]. [With emphasis:] Someone needs to keep the distance. And it should be us. (…) It's like a war, the generals have to decide where the bombs will fall, and they should not see the mess it creates. (Committee member 2, interview)

> [A good decision] requires a kind of distance from that specific [patient perspective]. (Committee member 3, interview)

To function well, the committee members feel they require metaphorical distance from the patients, which explains the lack of explicit discussion of argumentations contributed by patients or their representatives. Another committee member reflected:

> The patients challenge the committee to keep their position. (…) You need to stay detached. [But the patients' input] gives handles for substantiating [your position]: you must explain it well. It challenges you as a group and as a person. (Committee member 7, personal communication)

This committee member suggested that the patients' input increases the quality of the final (advised) decision as the process is "challenged" by the contributions.

Summarising, necessity is constructed by the committee during the deliberations by not weighing the patient (representative) contributions explicitly but by allowing them to challenge the decision-making process implicitly, which committee members consider heightens the quality of the justification or rationale for the decision.

Phase 3: deliberation

The next phase we describe is the deliberative discussion by the committee, where many different argumentations are contributed by the committee.

The discussion was usually initiated by the chair, with ample opportunity to speak and to respond to one another for the other committee members. The professed goal of the discussion is to ascertain whether there may be reasons to deviate from the reference value for cost-effectiveness (Committee member 5, personal communications), which range in three classes from 10 to 80,000 euro per Quality-adjusted Life Year (QALY). If the cost-effectiveness falls within a certain reference value range given a certain individual severity of illness, it is classed as favourable; if it does not, it is classed as unfavourable (Zorginstituut Nederland, 2017, 2018). The discussions both started and ended with a 'round' around the meeting table, where committee members were invited to speak in turn. In the deliberative phase of the meeting, members may make statements and respond to one another before the final decision (observations 1, 2, 4/audio files 1-4). The deliberations were subsequently summarised by the chair of the committee, based on which the secretary wrote the appraisal report, a summative report of the appraisal committee argumentations. The final advised decision was taken based on this combined document by the Institute's Board of Directors and sent to the Minister of Health who took the final decision (Couwenbergh et al., 2013; Zorginstituut Nederland, 2016, 2017).

The data on how necessity is constructed during the deliberations can be best characterised as *bringing new argumentations together*. This is visualised for the front teeth case, where the appraisal committee wrestled with the tension between not covering cosmetic (that is, as falling within the Range of Normality and thus unnecessary) surgery and making sure young people are able to do things like eating an apple (audio file 2, code: Societal Functioning).

This tension was resolved by one committee member, who humorously brought the following new necessity argumentations together:

> I had a strange association with the contraception debate, where we said, "You should pay for that yourself", but up to a certain age, we think that it needs to be reimbursed because of the situation that, just for example, a 14, 15-year-old with parents who think otherwise would not be able to – that it could result in unwanted pregnancies, and we would like 18-year-old girls to enter adulthood without an unwanted pregnancy. [Laughter] They must both have good teeth and not have an unwanted pregnancy! [Laughter]. (Committee member 9, observations/audio file 2, codes: Similar Treatments and Vulnerability/Compassion)

Bringing in new necessity argumentations and combining them was also visible in the maternity care case. The public regarded the preparation of *beschuit met muisjes* or *beschuitjes* (a traditional Dutch treat for friends and family paying maternity visits) by maternity care workers as unnecessary. The appraisal committee, however, did not consider this a decisive reason for denying coverage. These different viewpoints were expertly brought in and combined in the following way:

> If you look at that article in Trouw [Dutch newspaper] of this week, following the draft advice that was released, you'll see that maternity care is associated straightaway with prepping beschuitjes. [Laughter] And if at that point someone would say, "Wait a minute, er, should I pay for that?" I would have some sympathy with that. (...) [But] I think it's important to indicate something like: "Yes, but wait a minute, maternity care is about other matters, er, breastfeeding, detecting risky situations, etc., etc., for which it is completely just to be calling for solidarity". (Committee member 8, audio file 3, codes: Societal Responsibility and Vulnerability/Compassion)

Committee members and Institute employees describe the appraisal as an "open, moral" place "with permeable borders" where many "things" interact "organically" to form an advised decision (Committee members 2, 6, 4, Institute employee 4, (group) interviews). Specifically, the appraisal committee "brings in" new necessity argumentations "from the outside" to be "woven together" (Committee member 1, interview). In fact, this bringing in from the outside is part of their official task (art. 14, Zorginstituut Nederland, 2016).

In sum, necessity is constructed in this phase by bringing in new argumentations derived from the outside, from society, with sources including newspapers and previously-taken decisions, and weaving these considerations together.

Phase 4: decision

In the final phase of the appraisal, the committee formulates its positive or negative coverage decision advice.

The committee regularly gave additional recommendations, generally phrasing advice as: "yes, provided that…" or "no, unless…" (observations 1, 2, 4/audio files 1-4). For the eculizumab case and maternity care case, the advised decision was positive, provided that the work on the indication

protocols would continue (documents 1.2 and 3.2). Similarly, the paracetamol-vitamin D tablets are not covered except for certain medical indications; one of the committee members even summarised their deliberations as if directly giving the Minister advice. The primary rationale was that reimbursing these medicines would actually make them more expensive, due to the prescription rule:

> You may make many more [medicines] available outside the pharmacy. Given the situation, this is our answer: if it has to be bought at the pharmacy, it has to be reimbursed. But we advise you to think carefully about the prescription rule, because that creates a completely unequal ratio between those cheap medicines that are and those that are not available on prescription. (Committee member 6, audio file 4, code: Similar Treatments)

The committee thus gives recommendations to a broad set of actors including the manufacturer, the professional organisations involved in indication protocols, and the Minister of Health. We analyse this dynamic as a way of completing the construction of necessity: with these recommendations, the decision is placed back into society as it is linked directly not just to patients and the Minister of Health but to other stakeholders who will impact what care entails in practice.

DISCUSSION

In order to describe how necessity is constructed in Dutch health care coverage decisions, we followed the use of necessity argumentations across documents and the meetings of the appraisal committee at the Dutch Health Care Institute. Necessity is constructed by first, allowing explicit criteria contributed by the Institute employee to steer the process. Second, by allowing patient (representative) contributions to challenge the process: the decision should be sensitive to but not captured by particular interests. Notably, both the knowledge that is contributed by the Institute employee and the patients and/or their representatives implicitly shape the deliberations that follow. The third element we identify is bringing in new argumentations from the outside and weaving them together carefully. Fourthly and finally, necessity is constructed through formulating recommendations, making the decision more societally embedded than a tersely formulated 'yes' or 'no'.

In the introduction, we outlined our interest in societal weighing expertise and concluded that expertise in health care coverage decisions generally comprises 1) understanding different types of knowledge and 2) combining them into one decision, whilst 3) adhering not only to substantive requirements but also processual ones (Calnan et al., 2017; Gkeredakis et al., 2011; Hughes & Doheny, 2011; Jenkings & Barber, 2004; Russell, 2017; Russell & Greenhalgh, 2014). In terms of expertise pertaining to societal weighing specifically, we see that all three elements are confirmed by our dataset on constructing necessity to a certain extent. The committee indeed deals expertly with different types of knowledge. In this dataset these contain on the one hand argumentations

representing scientific knowledge contributed by the Institute employee (individual cost and severity of illness) but on the other hand also the patient (representative)'s experiential knowledge. We show that although these considerations are often not mentioned explicitly by the committee, they sometimes get rephrased, and they always steer the discussions implicitly, and are therefore considered crucial to the final decision. The committee also combines many argumentations and does so expertly. The overlap with coverage decision expertise, we argue, indicates that these parts of constructing necessity may indeed be classified as elements of societal weighing expertise and highlight the precise way that processual requirements, such as the presence of stakeholders, may impact appraisal deliberations. This relationship between necessity construction and expertise in societal weighing specifically is even stronger for the latter two phases of constructing necessity, as they show how argumentations are brought in from society (in our dataset, sources included newspapers and stakeholders), and the (advised) decision is embedded in society in turn. These two elements give a distinct flavour to societal weighing expertise that other studies of health care coverage decision-making expertise appear not to have hit upon.

The reluctance in terms of explicitly weighing the experiences of individual patients has previously been described for a variety of settings (Carlsen & Norheim, 2005; Hashem et al., 2018; Rooshenas et al., 2015). One potential underlying reason may be what Moreira describes as 'the politics of singularities'. Personal stories, according to Moreira, have a strong allegorical character by which they may spark the imagination through being relatable, and are thus able to destabilise other argumentations (Moreira, 2012). This fact that the committee listens to but does not explicitly mention these argumentations may be a manifestation of a refusal to be drawn into such politics. Regardless of the underlying reason, this finding is fascinating in light of recent widespread attempts to draw patients and citizens into such decision-making processes (Mitton et al., 2009; Wait & Nolte, 2006), also termed a 'multi-stakeholder appraisal' (Abrishami et al., 2017). Our data underline that this will not be easily achieved, which is in line with earlier work on the practices of dealing with different types of knowledge in health care coverage settings (Hashem et al., 2018; Moes et al., 2016).

On the recommendations specifically, the brunt of the available literature covers the process of coming to these decisions and the rationales behind them, rather than looking at what the additional recommendations might be (cf., Bukachi et al., 2014; Byskov et al., 2014; Giacomini et al., 2000; Madden et al., 2005; Martin et al., 2001; Rooshenas et al., 2015; Singer et al., 2000). Follow-up research may address questions on whether other appraisal committees also give recommendations, on the underlying dynamic these recommendations may point to, and implications for the process of health care coverage decisions generally.

Strengths and limitations

This chapter describes how a coverage appraisal is performed by expert decision makers and the dynamics of using argumentations therein. To our knowledge, this study adds to existing research on coverage decisions both methodologically, through showcasing how insight can be generated by

tracing argumentation types across documents and deliberative settings, as well as content-wise, noting what societal weighing expertise entails specifically. It makes a contribution to the literature of elegant muddling through by showing the emergent shared systematics behind it (Calnan et al., 2017; Russell, 2017). Moreover, it shows how pragmatic rationality is accomplished collectively; it is not just the committee but also the Institute employees and the patients that crucially shape the deliberations.

The methodology chosen will have impacted the data; the data set comprises a mixture of nine individual and three group interviews. Though both types of interviews were held primarily for purposes of data triangulation, the data gathered in these settings will have differed. In group interviews, the members of the group may stimulate each other (rather than the researcher being the only one to take this role) in terms of encouraging recall, opinion elaboration, and variation in response. However, group members may also correct each other, and even sway each other's opinions. Influential herein are group size, familiarity, and power dynamics (Frey & Fontana, 1991; King et al., 2018). In this data set, two of the three group interviews were with two direct Institute employee colleagues, who seemed high on familiarity and relatively low on power dynamics, positively impacting the data gathered. Interviewees did indeed often supplement each other; both double interviews were in fact suggested by the interviewees themselves for that reason. The third group interview, with four appraisal committee members, also concerned peers who were comfortable expressing their opinions together. Moreover, in this case, the quieter respondents were interviewed separately as well.

A major limitation of this study is the focus on four cases. It is relatively common to only focus on one case for characterising these types of decision making (Moes et al., 2016; Moreira, 2011). The case approach has granted us increased reliability but may necessarily lack some in-depth acquaintance with each case. Another limitation is the narrow focus on the deliberations in the appraisal committee meeting only. Other studies focus on the 'back stage', thereby uncovering more work that is done 'behind the scenes' to accomplish these types of deliberations (e.g., Escobar, 2015). Future research could attempt to visualise both, especially elucidating how the two intermingle in practice (cf. Hajer, 2005).

CONCLUSION

Using heterogeneous argumentations to make well-justified decisions is a task that many public institutions work hard to complete astutely. This chapter gives insight into the processes of tackling this task in a particularly vibrant field: health care coverage. It does so through examining the construction of necessity in the deliberative appraisal of four Dutch coverage decisions by following the necessity argumentations as mentioned by the different parties involved, supplemented by interviews with both appraisal committee members and Institute employees. Necessity is constructed differently in the four phases of the appraisal meeting, which, we show, correspond to four ele-

ments of expertise in societal weighing specifically. These elements comprise first, allowing explicit criteria to steer the process. Second, being shaped by the input of patient (representative)s; these are considered to challenge the process and heighten the quality of the justification or rationale. Third, bringing in new argumentations from society and weaving them together, and fourth, formulating recommendations to place the decision 'back' into society. These elements of societal weighing expertise explicate how the committee reaches a decision that is well-embedded in society.

REFERENCES

Abrishami, P., Boer, A., & Horstman, K. (2017). 4 Value in Co-Creation: Subjecting Innovative in-Hospital Technologies to Multi-Stakeholder Appraisal. *Public Value of Medical Innovations, 1*, 97.

Asdal, K. (2012). Contexts in Action—and the Future of the Past in Sts. *Science, Technology, & Human Values, 37*(4), 379-403.

Asdal, K., & Moser, I. (2012). Experiments in Context and Contexting. *Science, Technology, & Human Values, 37*(4), 291-306.

Berkhout, K. (2017). Slikken of Stikken? Het Kan Ook Anders, *NRC Handelsblad*.

Bukachi, S. A., Onyango-Ouma, W., Siso, J. M., Nyamongo, I. K., Mutai, J. K., Hurtig, A. K., . . . Byskov, J. (2014). Healthcare Priority Setting in Kenya: A Gap Analysis Applying the Accountability for Reasonableness Framework. *The International Journal of Health Planning and Management, 29*(4), 342-361. doi: doi:10.1002/hpm.2197

Byskov, J., Marchal, B., Maluka, S., Zulu, J. M., Bukachi, S. A., Hurtig, A.-K., . . . Consortium, T. R. (2014). The Accountability for Reasonableness Approach to Guide Priority Setting in Health Systems within Limited Resources – Findings from Action Research at District Level in Kenya, Tanzania, and Zambia. *Health Research Policy and Systems, 12*(1), 49. doi: 10.1186/1478-4505-12-49

Calnan, M., Hashem, F., & Brown, P. (2017). Still Elegantly Muddling Through? Nice and Uncertainty in Decision Making About the Rationing of Expensive Medicines in England. *International Journal of Health Services, 47*(3), 571-594.

Carlsen, B., & Norheim, O. F. (2005). "Saying No Is No Easy Matter" a Qualitative Study of Competing Concerns in Rationing Decisions in General Practice. *BMC health services research, 5*(1), 70. doi: 10.1186/1472-6963-5-70

Cerri, K. H., Knapp, M., & Fernandez, J.-L. (2014). Decision Making by Nice: Examining the Influences of Evidence, Process and Context. *Health Economics, Policy and Law, 9*(2), 119-141.

Commissie-Dunning. (1991). Kiezen En Delen. *Rapport van de Commissie Keuzen in de Zorg*.

Couwenbergh, B., Van Der Meer, F., Weghaus-Reus, S., Schelleman, H., & Zwaap, J. (2013). Pakketbeheer in De Praktijk Deel 3. *Diemen: Zorginstituut Nederland*, 1-106.

Creswell, J. W., & Poth, C. N. (2017). *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. USA: Sage publications.

Culyer, A. J., & Rawlins, M. D. (2004). National Institute for Clinical Excellence and Its Value Judgements. *British Medical Journal (BMJ), 329*, 224-227.

Dakin, H., Devlin, N., Feng, Y., Rice, N., O'neill, P., & Parkin, D. (2015). The Influence of Cost-Effectiveness and Other Factors on Nice Decisions. *Health Economics, 24*(10), 1256-1271. doi: doi:10.1002/hec.3086

Daniels, N. (2000). Accountability for Reasonableness. *Establishing a fair process for priority setting is easier than agreeing on principles, 321*(7272), 1300-1301. doi: 10.1136/bmj.321.7272.1300

Daniels, N., & Sabin, J. (1997). Limits to Health Care: Fair Procedures, Democratic Deliberation, and the Legitimacy Problem for Insurers. *Philosophy & Public Affairs, 26*(4), 303-350. doi: 10.1111/j.1088-4963.1997.tb00082.x

Daniels, N., & Sabin, J. (1998). The Ethics of Accountability in Managed Care Reform. *Health Affairs, 17*(5), 50-64. doi: 10.1377/hlthaff.17.5.50

Daniels, N., & Sabin, J. (2008). Accountability for Reasonableness: An Update. *Bmj, 337*. doi: 10.1136/bmj.a1850

Dussauge, I., Helgesson, C.-F., & Lee, F. (2015). *Value Practices in the Life Sciences and Medicine*: Oxford University Press, USA.

Escobar, O. (2015). Scripting Deliberative Policy-Making: Dramaturgic Policy Analysis and Engagement Know-How. *Journal of Comparative Policy Analysis: Research and Practice, 17*(3), 269-285.

Franken, M., Le Polain, M., Cleemput, I., & Koopmanschap, M. (2012). Similarities and Differences between Five European Drug Reimbursement Systems. *International Journal of Technology Assessment in Health Care, 28*(4), 349-357. doi: 10.1017/S0266462312000530

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, 119*(2), 195-202. doi: https://doi.org/10.1016/j.healthpol.2014.10.007

Frey, J. H., & Fontana, A. (1991). The Group Interview in Social Research. *The Social Science Journal, 28*(2), 175-187.

Giacomini, M., Hurley, J., & Stoddart, G. (2000). The Many Meanings of Deinsuring a Health Service: The Case of in Vitro Fertilization in Ontario. *Social Science & Medicine, 50*(10), 1485-1500. doi: https://doi.org/10.1016/S0277-9536(99)00394-9

Gkeredakis, E., Swan, J., Nicolini, D., & Scarbrough, H. (2011). *Rational Judgement Revisited: Practices of Deliberation in Healthcare Funding Decisions.* Paper presented at the OLKC conference, Hull University Business School.

Hajer, M. A. (2005). Setting the Stage: A Dramaturgy of Policy Deliberation. *Administration & society, 36*(6), 624-647.

Hashem, F., Calnan, M. W., & Brown, P. R. (2018). Decision Making in Nice Single Technological Appraisals: How Does Nice Incorporate Patient Perspectives? *Health Expectations, 21*(1), 128-137. doi: doi:10.1111/hex.12594

Hasman, A., & Holm, S. (2005). Accountability for Reasonableness: Opening the Black Box of Process. *Health Care Analysis, 13*(4), 261-273. doi: 10.1007/s10728-005-8124-2

Hoedemaekers, R., & Oortwijn, W. (2003). Problematic Notions in Dutch Health Care Package Decisions. *Health Care Analysis, 11*(4), 287-294. doi: 10.1023/B:HCAN.0000010057.43321.b2

Hughes, D., & Doheny, S. (2011). Deliberating Tarceva: A Case Study of How British Nhs Managers Decide Whether to Purchase a High-Cost Drug in the Shadow of Nice Guidance. *Social Science & Medicine, 73*(10), 1460-1468.

Jansen, M. P., Helderman, J.-K., Boer, B., & Baltussen, R. (2017). Fair Processes for Priority Setting: Putting Theory into Practice : Comment on "Expanded Hta: Enhancing Fairness and Legitimacy". *International Journal of Health Policy and Management, 6*(1), 43-47. doi: 10.15171/ijhpm.2016.85

Jenkings, K. N., & Barber, N. (2004). What Constitutes Evidence in Hospital New Drug Decision Making? *Social Science & Medicine, 58*(9), 1757-1766.

King, N., Horrocks, C., & Brooks, J. (2018). *Interviews in Qualitative Research*. UK: SAGE Publications Limited.

Kleijnen, S., George, E., Goulden, S., D'andon, A., Vitré, P., Osińska, B., . . . Nagy, B. Z. (2012). Relative Effectiveness Assessment of Pharmaceuticals: Similarities and Differences in 29 Jurisdictions. *Value in Health, 15*(6), 954-960.

Kleinhout-Vliek, T., De Bont, A., & Boer, B. (2017). The Bare Necessities? A Realist Review of Necessity Argumentations Used in Health Care Coverage Decisions. *Health Policy, 121*(7), 731-744. doi: https://doi.org/10.1016/j.healthpol.2017.04.011

Kleinhout-Vliek, T., de Bont, A., Boysen, M., Perleth, M., van der Veen, R., Zwaap, J., Boer, B. (2020). Around the Tables – Contextual Factors in Healthcare Coverage Decisions Across Western Europe. International Journal of Health Policy and Management, 9(9), 390-402. doi: 10.15171/ijhpm.2019.145

Lamont, M. (2012). Toward a Comparative Sociology of Valuation and Evaluation. *Annual Review of Sociology, 38*(1), 201-221. doi: 10.1146/annurev-soc-070308-120022

Le Polain, M., Franken, M., Koopmanschap, M., & Cleemput, I. (2010). Drug Reimbursement Systems: International Comparison and Policy Recommendations. *Health Services Research (HSR), KCE Reports C, 147.*

Madden, S., Martin, D. K., Downey, S., & Singer, P. A. (2005). Hospital Priority Setting with an Appeals Process: A Qualitative Case Study and Evaluation. *Health Policy, 73*(1), 10-20. doi: https://doi.org/10.1016/j.healthpol.2004.11.002

Makady, A., Ten Ham, R., De Boer, A., Hillege, H., Klungel, O., & Goettsch, W. (2017). Policies for Use of Real-World Data in Health Technology Assessment (Hta): A Comparative Study of Six Hta Agencies. *Value in Health, 20*(4), 520-532.

Martin, D., Pater, J., & Singer, P. (2001). Priority-Setting Decisions for New Cancer Drugs: A

Qualitative Case Study. *The Lancet, 358*(9294), 1676-1681. doi: https://doi.org/10.1016/S0140-6736(01)06714-9

Mastenbroek, C., Van Der Meer, F., Zwaap, J., Rikken, F., & Polman, P. (2006). Pakketbeheer in De Praktijk. *Diemen: College voor zorgverzekeringen*.

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.

Moes, F., Houwaart, E., Delnoij, D., & Horstman, K. (2016). Contested Evidence: A Dutch Reimbursement Decision Taken to Court. *Health Economics, Policy and Law, 12*(3), 325-344. doi: 10.1017/S1744133116000281

Moreira, T. (2011). Health Care Rationing in an Age of Uncertainty: A Conceptual Model. *Social Science & Medicine, 72*(8), 1333-1341.

Moreira, T. (2012). Health Care Standards and the Politics of Singularities:Shifting in and out of Context. *Science, Technology, & Human Values, 37*(4), 307-331. doi: 10.1177/0162243911414921

Nice. (2008). *Social Value Judgements - Principles for the Development of Nice Guidance*. https://www.nice.org.uk/Media/Default/about/what-we-do/research-and-development/Social-Value-Judgements-principles-for-the-development-of-NICE-guidance.pdf.

Niëns, L. (2014). Affordability in Health Care: Operationalizations and Applications in Different Contexts.

Oliver, A., Mossialos, E., & Robinson, R. (2004). Health Technology Assessment and Its Influence on Health-Care Priority Setting. *International Journal of Technology Assessment in Health Care, 20*(1), 1-10. doi: 10.1017/S026646230400073X

Patera, N., & Wild, C. (2014). Assessment – Appraisal – Decision. Lbi-Hta Decision Support Document Nr.: 72. Wien: Ludwig Boltzmann Institut für Health Technology Assessment.

Poley, M. J., Stolk, E.A., Brouwer, W. B. F., Van Busschbach,J. J. . (2002). Ziektelast Als Uitwerking Van Het Criterium 'Noodzakelijkheid' Bij Het Makenvan Keuzen in De Zorg. *Ned Tijdschr Geneeskd, 146*(4), 2312-2315.

Ragin, C. C. (2004). Turning the Tables: How Case-Oriented Research Challenges. *Rethinking social inquiry: Diverse tools, shared standards, 123*.

Reckers-Droog, V., Van Exel, J., & Brouwer, W. (2018). Looking Back and Moving Forward. *Health Policy*.

Reckers-Droog, V., Van Exel, J., & Brouwer, W. (2019). Equity Weights for Priority Setting in Healthcare: Severity, Age, or Both? *Value in Health, 22*(12), 1441-1449.

Rooshenas, L., Owen-Smith, A., Hollingworth, W., Badrinath, P., Beynon, C., & Donovan, J. L. (2015). "I Won't Call It Rationing…": An Ethnographic Study of Healthcare Disinvestment in Theory and Practice. *Social Science & Medicine, 128*(Supplement C), 273-281. doi: https://doi.org/10.1016/j.socscimed.2015.01.020

Russell, J. (2017). *The Rationality of Rationing: A Rhetorical Policy Analysis of Deliberations About Resource Allocation in the Nhs*. University of Oxford, UK.

Russell, J., & Greenhalgh, T. (2014). Being 'Rational'and Being 'Human': How National Health Service Rationing Decisions Are Constructed as Rational by Resource Allocation Panels. *Health:, 18*(5), 441-457.

Salas-Vega, S., Bertling, A., & Mossialos, E. (2016). A Comparative Study of Drug Listing Recommendations and the Decision-Making Process in Australia, the Netherlands, Sweden, and the Uk. *Health Policy, 120*(10), 1104-1114.

Samenleving, R. V. V. E. (2017). Zonder Context Geen Bewijs. *Over de illusie van evidencebased practice in de zorg. Den Haag: RVS*, 17-05.

Sayer, A. (2011). *Why Things Matter to People: Social Science, Values and Ethical Life*: Cambridge University Press.

Shah, K. K., Cookson, R., Culyer, A. J., & Littlejohns, P. (2013). Nice's Social Value Judgements About Equity in Health and Health Care. *Health Economics, Policy and Law, 8*(2), 145-165.

Shirazi, P. A., Boer, A., & Horstman, K. (2017). Value in Co-Creation: Subjecting Innovative in-Hospital Technologies to Multi-Stakeholder Appraisal. *International Journal of Hospital Based Health Technology Assessment (IJHBHTA)*, 12-30.

Singer, P. A., Martin, D., Giacomini Bhatia, M. V., & Purdy, L. (2000). Priority Setting for New Technologies in Medicine: Qualitative Case Study. *Bmj, 321*(7272), 1316-1318. doi: 10.1136/bmj.321.7272.1316

Stolk, E., Brouwer, W., & Busschbach, J. (2002). Rationalising Rationing: Economic and Other Considerations in the Debate About Funding of Viagra. *Health Policy, 59*(1), 53-63.

Stolk, E., Goes, E., Kok, E., & Busschbach, J. (2001). Uitwerking Criteria Noodzakelijkheid, Eigen Rekening En Verantwoording En Lifestyle. *College voor Zorgverzekeringen Breedte geneesmiddelenpakket. Amstelveen, The Netherlands: College voor Zorgverzekeringen, 1*, 54.

Unknown. (2015). Kraamzorg Hoort in Het Basispakket, *Trouw*.

Unknown. (2016). Zware Paracetamol Mogelijk Uit Basispakket: Klap Voor Honderdduizenden Patiënten, *RTL nieuws*.

Van Der Aa, E. (2016). Zware Paracetamol En Vitaminen Wellicht Uit Pakket, *Algemeen Dagblad*.

Vuorenkoski, L., Toiviainen, H., & Hemminki, E. (2008). Decision-Making in Priority Setting for Medicines—a Review of Empirical Studies. *Health Policy, 86*(1), 1-9.

Wait, S., & Nolte, E. (2006). Public Involvement Policies in Health: Exploring Their Conceptual Basis. *Health Economics, Policy and Law, 1*(2), 149-162.

Walley, T. (2007). Health Technology Assessment in England: Assessment and Appraisal. *Medical Journal of Australia, 187*(5), 283.

Zorginstituut Nederland. (2016). Reglement Adviescommissie Pakket. In Z. Nederland (Ed.). https://www.zorginstituutnederland.nl/over-ons/publicaties/besluit/2016/12/05/reglement-adviescommissie-pakket-zorginstituut-nederland (accessed 11 February 2020).

Zorginstituut Nederland. (2017). Pakketadvies in De Praktijk: Wikken En Wegen Voor Een Rechtvaardig Pakket. Diemen.

Zorginstituut Nederland. (2018). Ziektelast in De Praktijk. Diemen.

4 | Conceptual model for robust decisions

Modified version of paper under review at Sociology of Health & Illness as: *Kleinhout-Vliek, T., de Bont, A., & Boer, B. Arguments in Networks – a conceptual model for robust health care coverage decisions*

INTRODUCTION

Long-debated and carefully made negative health care coverage decisions are sometimes followed by extensive media coverage and public discussion. Health care coverage, which deals with whether a health care technology is provided or not on a national level, is a contentious subject when the headlines, featuring sick people no longer receiving their too-expensive medicines, virtually write themselves (Abelson & Collins, 2009). Not rarely, public controversy results in the reversal of the original decision. These reversals often happen through a direct appeal to the Minister of Health or other public authority, circumventing the original decision process and the reasons given therein (Aggarwal et al., 2014; Burls et al., 2005; Clarke et al., 2001; MacKenzie et al., 2008; McIver & Ham, 2000). Negative coverage decisions may also be appealed in court (Moes et al., 2017; Moreira, 2011). Such court proceedings may likewise circumvent the original rationale, even when they do not overturn the original decision. They deal instead with whether due procedure was followed in e.g., coming to the final decision (Moreira, 2011) or in establishing the knowledge upon which the decision was made (Moes et al., 2017). The public controversy around these coverage decisions show that the reasons provided for the original decision were considered inappropriate or inadequate.

Decisions so controversial that they are circumvented do remain the exception: the majority of decisions pass relatively unnoticed by the public eye, the reasons provided remain unquestioned; they cannot be considered controversial in that sense. These decisions do, however, need resolve the perceived tension between the individual and the collective good (cf. Prainsack, 2018). This tension is inherent to all coverage decisions, embodied respectively by on the one hand, the general rules or criteria laid out for such decisions and, on the other hand, the patients present in the deliberative setting (Carlsen & Norheim, 2005; Hashem et al., 2018; Moreira, 2011, 2012; Rooshenas et al., 2015). Resolving this perceived tension is generally achieved through finding pragmatic ways to come to a rationale (Callon et al., 2009; Calnan et al., 2017; Moreira, 2011; Russell & Greenhalgh, 2014). As shown above, public controversies have proven fruitful research sites in terms of showing forms of reasoning that may otherwise be taken for granted (Moreira, 2011). Little is known about how rationality is constructed, how the inherent tension is dealt with, in taken-for-granted ways (cf. Lehoux et al., 2010).

This chapter will address the question: how are relatively robust run-of-the-mill coverage decisions made? We will draw from Science and Technology Studies (STS) literatures on controversy and robustness. We will conceptualise robust decision outcomes, including the publicly available reasons provided, as networks of elements that are able to withstand pressure 'out there', and describe how this is achieved through three steps: identification of elements, designing networks of elements, and through testing these differing networks. This will be illustrated by data gathered in previous qualitative comparative case study research in the setting of Dutch health care coverage decisions and will lead to several recommendations for policy and research.

ROBUST DECISIONS AS NETWORKS

Rip (1985, 1986, 2010) embarks upon his operationalisation of the robustness of decisions through what he terms an 'informal technology assessment', in essence a public litmus test for decision outcomes. Robust outcomes, Rip writes, are able to withstand "the pressures to which they will inevitably be exposed" (Moreira, 2013; Rip, 1985). These public pressures thus test robustness: the proof is in the pudding. The STS field may pride itself on a long history of studying this pudding, namely public controversies, as fruitful sites for exploration of the role of technology in society (Jasanoff, 2012; B. Martin & Richards, 1995). Scholars describe how polities have dealt with controversies as diverse as nuclear power plants, radio-active waste storage, Bovine Spongiform Encephalopathy, HIV/AIDS, Genetically Modified Organisms, nanotechnology, but also coverage decisions (Callon et al., 2009; Epstein, 2011; Moreira, 2011). These studies describe the work that is put in to uphold or defuse a controversy, highlighting the insufficiency of traditional, 'certified' expertise, as always making the reader sensible to the many ways in which assemblages of the elements of the controversy come to be, and in which closure was perhaps achieved (B. Martin & Richards, 1995). However, as Rip contends: "the analyst of controversies should limit himself to identify robustness and trace the processes that produce it" (Rip, 1986). This shows an important distinction: many STS scholars study the 'public life' of controversies, whereas Rip's primary object is the robustness as the outcome.

Starting with Rip, then, we read that robust outcomes contain:

> Arguments, evidence, social alignments, interests, and cultural values, many of them interrelated and therefore lending support to the dominant view. (Rip, 1986, p. 353)

Elsewhere, Rip speaks of "alignment of findings, arguments, perceptions, interests, and dominant values – and circumstances." (Rip, 2010). From this, we derive a first important aspect of a robust outcome: robust outcomes are based on multifarious elements of different kinds. Rip continues,

> The difference between an only fashionable and a robust view is a matter of degree, and perhaps also a matter of actual effort that actors are prepared to exert. In both cases, the views are available in the cultural repertoire, but with increasing robustness, the linkages between elements of the view and with their context increase in number and in articulation (and sometimes also in scope). (Rip, 1986, p. 353)

Two more aspects of robust outcomes according to Rip are then, first, making these networks of elements requires effort as it is achieved actively, and this effort exceeds just bringing all these elements together. Setting up robust networks of elements involves matching up, 'clicking together' some of these elements (Rip, 1986). Second, Rip speaks of the characteristics of the links between elements: as they are articulated they increase in number, and sometimes in scope. What follows

logically is that the making of these links is an important, powerful act in achieving robustness. Rip distinguishes two separate linking activities, namely articulation and consolidation. By articulation, he understands that elements are actively joined together that were previously unlinked, as described above, including increasing the scope of the decision. Consolidation, for Rip, is the next step in terms of robustness, where several elements are so strongly linked together that the link itself is 'black-boxed' and becomes difficult to call into question (Rip, 1986).

Importantly, Rip highlights the availability of different possible networks of elements, which differ in robustness. Some are on the "fashionable" end of the spectrum, while others are more robust. This theme has been expanded by Callon, Lascoumes and Barthe, who highlight the productivity of controversy. For them, controversy yields a robust solution because it "allows the *design* and *testing* of (…) [multiple] solutions that integrate a plurality of points of view, demands, and expectations" (Callon et al., 2009, p. 32, emphasis added). The idea that controversy is not just a fruitful research site but that controversies may also be productive in a broader sense has also been described by Rip, who already noted in 1986 that governments may have a specific task to encourage learning from controversy (Rip, 1986); Nowotny even conceptualises robustness as resulting specifically from such repeated testing (Nowotny, 2003). Callon et al specify how this learning-from-controversy should be encouraged: namely through what they term *hybrid forums*. In these, specialists and laypersons design and test plural solutions to the controversy, or networks of elements, together (Callon et al., 2009). In this, the designing resonates clearly with the work of Rip as described above. The testing process, aimed at "taking more into account", Callon et al. specify to entail a series of negotiations and compromises between all present (Callon et al., 2009, p. 32). Callon et al. argue in sum that controversy should be made part and parcel of the decision-making process instead of remaining 'out there': controversy is to be actively encouraged in a 'safe' space as more, and more different, people have a stake and a say.

Health care coverage offers a vibrant topic to explore the productivity of controversy in such hybrid forums (Moreira, 2013). Moreira argues that hybrid forums dealing feature 'hybrid knowledge' (cf. Nowotny, 2003), a form of understanding that exceeds the purely technical and involves the exploration of links between different arguments (Moreira, 2011). Such an exploration of links strongly resembles what Callon et al. call the design of a solution, and what is visible in Rip's work as the process of bringing a decision outcome together; what we will conceptualise as designing linked networks of elements. This designing step takes place after the identification of the elements as proposed by Rip, which is therefore the first step of our model. These elements here comprise of findings (that is, evidence that is considered 'scientific'), arguments, and values – in the knowledge that this is a reduction of the variety of elements present. These three types are generally present in health care coverage decisions: scientific findings of experts in the shape of Health Technology Assessments and the like; arguments such as concerning what is considered good care; and values such as justice, equality, and solidarity (Calnan et al., 2017; Lehoux, 2014; Lehoux et al., 2010; Makady et al., 2017; Vuorenkoski et al., 2008). Step 2 of our proposed model then, designing the networks, is specifically conceptualised as forging links between all these different elements.

These links are actively articulated, increased in scope, and sometimes even 'black-boxed'. After the elements have been identified and the networks designed, the third step is testing the multiple solutions, that is, the decision networks – and choosing one.

In the remnant of this chapter we will draw on several years of field work on the use of argumentations in the appraisal of health care coverage technologies, to show how the decision trajectories of two specific cases illustrate this 3-step model as they link multiple, diverse elements into a network which is relatively robust.

COMPARATIVE CASE STUDY METHODOLOGY

This chapter builds on previous research at the Dutch Health Care Institute, which employed a case studies approach (Creswell & Poth, 2017; Ragin, 2004). Case study analysis is well-placed to provide insight into health care coverage decisions as it gives an in-depth take on processes that entail valuation (Dussauge et al., 2015; Lamont, 2012). We opted for two cases with high contrast (Lamont & Thévenot, 2000) in terms of decision outcome, number of patients affected, type of technology, and price: eculizumab (Soliris®) and paracetamol-vitamin D tablets (for more information, see boxes 1 and 2). Both were discussed in the Dutch media but caused little to no public controversy (Berkhout, 2017; Unknown, 2016; Van der Aa, 2016).

The data consisted of observations at appraisal committee meetings (n=3, one meeting for eculizumab and two for paracetamol-vitamin D tablets, with field notes taken), interviews with committee members (n=4 in three interviews) and with Institute employees (n=2 in one interview). The interviews were all semi-structured, and the two group interviews were of the field-formal type, both with the purposes of data triangulation (Frey & Fontana, 1991; King et al., 2018). Moreover, five documents pertaining to the two decisions were analysed, and the audio files for the appraisal committee meetings obtained to reference the field notes. The secretary of the appraisal committee granted access to the meetings and (audio) files. Data analysis and formulation of the conceptual model was enriched by other previous observations at appraisal committee meetings (n=4) and interviews with committee members (n=6) and Institute employees (n=7 in six interviews) (see Tables 1 and 2).

Case study	Documents and audio files
Eculizumab (Soliris®)	discussion document appraisal report patient contribution (part of discussion document) audio file
Paracetamol-vitamin D tablets	discussion document appraisal report audio file

Table 1 - overview of documents and audio files analysed per case

Interviews and observations date	Description
November 2015	Observations appraisal committee paracetamol-vitamin D tablets
October 2016	Committee member 5
October 2016	Institute employees 1 & 7
October 2016	Observations appraisal committee eculizumab and paracetamol-vitamin D
February 2017	Committee members 1, 4, 5 & 6
October 2017	Committee member 6

Table 2 - overview of interviews and observations analysed

DUTCH HEALTH CARE COVERAGE DECISIONS

In the Netherlands, all citizens are insured through private insurance, which covers at least the bare minimum set by the government: the basic benefits basket. The Dutch National Health Care Institute (in Dutch: *Zorginstituut Nederland*, in the rest of this chapter: 'the Institute') is responsible for advice to the Minister of Health as to the contents of this benefits basket. The Institute utilises four formalised criteria to come to this advice. These are: 1) effectiveness and 2) cost-effectiveness of the health care technology, 3) feasibility of coverage (including total budget impact), and 4) necessity, which rests in part on the individual severity of illness and individual affordability (Zorginstituut Nederland, 2013). These well-established criteria (Franken et al., 2012) are brought together in the final advice document, which is sent to the Minister, who makes the final decision.

The process of coming to this advice comprises several steps, relatively common in its set-up (Patera & Wild, 2014; Walley, 2007). Generally, it starts with a scoping session in which stakeholders are invited to submit initial comments on the health care technology under consideration. This is followed by the assessment phase, wherein the scientific evidence, which includes information on the effectiveness, cost-effectiveness, budget impact, severity of illness, and affordability, prepared by Institute employees, is rigorously examined by the assessment committee (in Dutch: *Wetenschappelijke Adviesraad*). The outcomes of this examination are subsequently contextualised in the meeting of the appraisal committee (in Dutch: *Adviescommissie Pakket*). In this meeting, which specifically aims to 'bring in' wider societal values pertaining to the technology under consideration, the scientific evidence is first presented by an Institute employee, followed by input from patients (representative)s and/or other stakeholders. Subsequently, the appraisal committee, which comprises eight to ten external experts from fields such as pharmaco-economics and health care ethics, commences their deliberations. At the end of these deliberations, in which every member is expressly given an opportunity to speak and respond, the final advice is formulated. This advice is then summarised, approved by the Institute's Board of Directors, and forwarded to the Minister (Zorginstituut Nederland, 2016, 2017).

THREE-STEP MODEL FOR ROBUST DECISIONS

Our conceptual model comprises three steps towards robust decisions, which we conceptualise as networks of diverse elements. The first step is identification of many different elements (findings, arguments, values). Second, networks are designed from these elements through linking elements, broadening the scope of the network, and black-boxing links. For the third and final step, the networks that have thus been designed are tested for robustness, including the moment where one network is chosen over alternative networks.

Step 1: Identifying elements

The first step in coming to a robust decision is to identify different elements that may make up the network. We are not suggesting 'all' elements can be identified; rather, efforts in this area are rewarded with many different elements, and many different types of elements. One way to obtain many elements is through inviting experts-by-experience and laypeople, such as patients or other members of the public, into the deliberative setting of the decision. In the Dutch appraisal committee, deliberations do indeed start with contributions from Institute employee(s) and, sometimes, patient (representative)s. The appraisal actively invites these contributions, treating them as valuable, especially as structuring the deliberations in terms of coming to an advised decision (Cerri et al., 2014).

These contributions contain the three types of elements of networks: findings, arguments, and values. These three element types are clearly visible in the dataset on eculizumab (see Box 1) and paracetamol-vitamin D tablets (see Box 2). *Findings* included the individual severity of illness, contributed by the Institute employee and by patient representatives, and the extent to which the societal functioning of the patient is hindered by the disease, described by the patient representative in the case of eculizumab. *Arguments* include the small number of patients, making necessity of coverage greater because of the difficulty in obtaining trustworthy cost-effectiveness data, mentioned by the Institute employee in the case of eculizumab, and the absence of viable alternatives, remarked upon by both the Institute employee and a patient representative for eculizumab. For paracetamol-vitamin D tablets, arguments contributed by pharmacists included the fact they considered substitution to heavier medicines likely. *Values*, then, include the justice of coverage and a clear call for compassion, also contributed by a patient representative for eculizumab, In the case

Eculizumab (Soliris®) is an orphan drug licensed for Atypical Hemolytic Uremic Syndrome (aHUS) and Paroxysmal Nocturnal Hemoglobinuria (PNH). The cost-effectiveness ratio of the medicine was considered "highly unfavourable", usually a decisive criterion (as visible in many coverage decisions for cancer drugs, as well as the PNH decision, which was negative). Patients and clinicians had, however, taken the initiative to research methods to shorten the treatment period, expected to result in a more favourable cost-effectiveness ratio. The deliberations in the appraisal committee hinged on this independent research; they considered the burden of disease combined with the risk taken by the patients so commendable that it yielded positive advice, provided the work on the research protocol and indication criteria would continue (unpublished results).

Box 1 – eculizumab

The Institute was set to make an advised decision on a list of tablets, collectively named 'paracetamol-vitamin D', after it had been set on the agenda by the Minister of Health. This in response to changes on a list of covered medicines (*Geneesmiddelenvergoedingssysteem* in Dutch), on which these tablets were 'left behind'. The Minister of Health explicitly asked the Institute whether these were in fact "necessary to be insured". The final advised decision was to no longer cover these tablets, except for certain medical indications. The reason given was that covering these relatively inexpensive tablets would turn out to be more expensive, because of the 'prescription rule'. This rule is an additional charge levied by the pharmacist when a covered medicine is bought at the pharmacy on prescription (unpublished results).

Box 2 – paracetamol vitamin D

of paracetamol-vitamin D tablets, values mentioned by a committee member, speaking on behalf of patients, included solidarity with vulnerable groups (audio files #1 and #2/field notes #161014).

This overview showcases that the provenance of elements may be somewhat counter-intuitive. Findings, for example, are not only brought in by certified experts. Rather, individuals speak as a

> Collection of voices (…) sometimes speaking as a patient, as a doctor, as a member of the public, from the voice of medicine, the voice of the life world, speaking as both a panel 'insider' and 'outsider', sometimes even within the same utterance. (Russell, 2017, p. 219).

This is also visualised by a committee member in the eculizumab case, who stated that:

> You want this [initiative] to be rewarded! The cost-effectiveness improves highly significantly. (…) Postponing [the decision] would be bad for the patient and the [health insurance] premium payer. (…) Economically speaking, the logic is zero. (Committee member 11, audio file #1)

This committee member moves deftly and quite naturally from his or her personal response, to the criterion of cost-effectiveness, to potential effects on patients and members of the public, to an economic line of reasoning. Such dynamics raise questions on the identity and distinctiveness of expertise and its role in policy making – not for nothing has this been a long-standing debate in STS and adjoining fields (Callon, 1999; Collins et al., 2010; Fischer, 2011; Jasanoff, 2003; Nowotny et al., 2013; Rip, 1985, 2003). In terms of both these aspects, we would follow Callon et al. (2009) and Moreira (2011, 2013) in not focusing on teasing out the differences between contributions but holding that contributions of more types of people yields more types of elements. Naturally, these elements may clash with one another; some will argue in favour of, and others against coverage of this particular health care technology (Kleinhout-Vliek et al., 2017). This is not problematic, however, as they may simply become part of different networks (see step 2 below).

Step 2: Designing networks

In arriving at a robust decision, the second step is designing the networks of elements (Callon et al., 2009), where three activities may be separated out: articulating links between elements, broadening the scope of the network, and black-boxing links (Rip, 1986).

Articulating links is the primary method for connecting elements into decisions (D. Martin et al., 2001; Singer et al., 2000). The document with reasons provided for a decision always contains a variety of both formalised criteria and case-specific argumentations (Kleinhout-Vliek et al., 2020). Links are made all the time in deliberations, and we are specifically interested in links made between different types of elements. This is visible in the paracetamol-vitamin D case, where it was argued that:

> People who take these medicines often have more costs due to comorbidity and/or cannot afford them because of a low average income coupled to a lower socio-economic status. (Discussion document #2).

This is a clear linking of the value of equity with the finding of severity of illness and the argument of personal responsibility. For the eculizumab case, a strong linked network was already available, because a negative coverage decision had been made on the same medicine for a different indication, PNH (see Box 1). The new links articulated included the value placed on the courage of the patients in terms of shouldering the uncertainty of relapse, and the fact that they should not "be duped by the manufacturer's criminal acts [in terms of repeatedly failing to provide high-quality data]" (Committee member 10, audio file #1/field notes #161014).

Some of these elements, such as the negligence of the manufacturer in terms of providing evidence of sufficient standard had been noted before, though it had not impacted the PNH decision in the same way. This shows that though some elements had been contributed before, they may become part of *different* networks (cf. Callon et al., 2009). It thus benefits the deliberations that the elements and the subsequent links are made *in situ*: it highlights the expertise brought to bear in articulating these links between different elements.

Broadening the scope of the network appears to be not common in the published literature, where decisions hinge on explicated reasons and are rarely described to concern other areas of health care (cf. Rip, 1986). However, the coverage decisions studied contain recommendations, and we pose this falls squarely in this category. In the case of the paracetamol-vitamin D case study, one of the appraisal committee members broadened the scope of the coverage decision, formulating it as if directly giving the Minister of Health advice on the prescription rule (see Box 2). This is not part of the remit of the basic benefits basket, or indeed, of the appraisal committee. However, recommendations like these provide an essential strengthening element of a network: the scope of the decision is broadened by going beyond the coverage decision. Specifically, the direction in which the network is broadened through such recommendations remains at the committee's discretion. Sometimes they formulate advice to the Minister, sometimes to other stakeholders, such as committee member 13, who calls for the expertise centre to communicate to other countries the desire to pressurise the manufacturer (audio file #1/field notes #161014). In this way, the committee does not only specify what a good basic benefits basket is but actively broadens its remit to specifying what good *care* entails.

Black-boxing links is the final and most robust aspect of designing networks. The example provided by Rip concerns the black-boxed link between smoking and cancer (Rip, 1986, p. 354). This, even more than broadening the scope of the network, may be considered unchartered territory. Coming from a strong history of providing elaborate reasonings, with as many (predetermined) elements and links between elements visualised as possible (eg., Bærøe & Baltussen, 2014; Guindo et al., 2012), we argue the idea of black-boxing may seem counter-intuitive. However, cost-effectiveness in and of itself could be seen as black-boxed, linking many separate elements such as quality-adjusted life years, costs per treatment, and effectiveness, into a widespread coverage criterion (Franken et al., 2012). In this sense, the committee's deliberations (almost) always employ a link that has been black-boxed for a long time. Black-boxing is also visible to some extent in the appraisal committee's work we have studied, for example in the advice on eculizumab. The potential network of a negative decision, where the costs per QALY are insufficient given the severity of illness, is common, even so common that it has been described as "simply stamping the file" (Committee member 6, personal communications). The black-boxing here has happened through past decisions, in which that stamp was formed, allowing it to be "simply" used in later decisions. Other black-boxed links include another negative decision linking the finding of affordability (with the rule of thumb of cheaper than €100) and the argument of it being cheaper *not* to provide through the collective health insurance as this will keep prices up. Black-boxing links thus happens (but is naturally not explicated), with some links so black-boxed they will be difficult to prise open.

Step 3: Testing networks

The third and final step is to test these different designed decision networks. That different networks may exist and one needs to be chosen has previously been described for a decision for a highly expensive treatment, which gained a positive coverage status because one set of clustered argumentations together weighed more heavily than another set (D. Martin et al., 2001; Singer et al., 2000). Callon et al. describe this process as a series of negotiations and compromises in the hybrid forum setting, which "unleashes" the learning process to be harnessed as part of the decision (Callon et al., 2009, p. 32; Moreira, 2011). The confrontation between laypersons and those holding expertise is vital for enabling this learning process. This is illustrated by the eculizumab case, where, as stated above, one relatively strong network was already available. As committee member 2 summarised:

> Over the past few months, we have said no regularly, for other '–mabs' [a class of expensive drugs]. Not because we would begrudge patients their medicines but because we think prices should be reasonable. (Committee member 2, field notes #161014)

The network of this well-established negative decision was tested by this comment – all subsequent commenters contributed not to strengthening this network but to constructing the alternative. The first consideration that came up regularly was the courage of the patients and the uncertainty they had had to live with because no final coverage decision had been taken for four years. The

second was the high valuation of the proposed research protocol, especially as it was conducted independently from the manufacturer, which was considered likely to also positively affect the price (field notes #161014). We analyse the repetition as a testing of the alternative network: it becomes clear, as one committee member after the other contributes, that this is where all want to go. As committee member 6 stated, summarising:

> The rules should be stretched – we want to make something possible here. (Committee member 6, audio file #1/field notes #161014)

This example shows how committee members work together in testing the networks. In other cases, it may be done by one committee member. As visible in the paracetamol-vitamin D case excerpt below, the individual responsibility for the 'bottom' (cheaper end) of the benefits basket is noted, informally, to apply to any medicine cheaper than €100 per year. However, this particular decision required an opening up of what is usually black-boxed, because of the strong alternative network of vulnerable groups. The committee member does so through making *more* links to this network, and then, dismissing this network through noting it as another's responsibility.

> Through the argument of not being able to afford [the tablets] it seems as if (…) there is some sort of poverty boundary where people through the calcium tablets will suddenly end up on the wrong side. But in those cases, (…) there is probably already more going on, with those people, already the government, all kinds of related measures, rent subsidies, benefits, are happening. And it won't [mean] those benefits agencies will give extra benefits because of this tablet, but there is a whole host of expenses, gas, light, and oh yes, the costs of this medicine, so there's much more to it. I think it is almost, how should I say it, almost a self-centred idea that we or [the collective health] insurance were going to make the difference between poverty and no poverty. So I would argue to leave the poverty and the not being able to afford [things] to agencies that deal with these things, [because] it will not be influenced by that one-and-a-half calcium tablet (…). You can talk about those 100 euros but it is always, low cost – just put it aside. (Committee member 4, audio file #2).

First, this is a notable remark as it narrows the committee's responsibility, counter to broadening it (see above). Second, this network was quickly considered decisive; the committee responded primarily through noting that this problem indeed should not be solved through health insurance (Committee member 6, audio file #2) and the fact that "the whole system is inefficient" (Committee member 4, audio file #2). It was, however, source of little discussion. This shows that sometimes, testing a network is done by one committee member who then changes the direction of the whole discussion (as for the paracetamol-vitamin D case), but sometimes, it is a collaborative effort, which is initiated tentatively and resonates around the room (as in the eculizumab case). As committee member 6 said, reflecting on the paracetamol-vitamin D case:

We just cut the Gordian knot on very pragmatic grounds. (Committee member 6, interview)

Later, in personal communications, the same committee member said that s/he considered the eculizumab advised decision to be one of the most beautiful ones s/he had contributed to. We would analyse this committee member to be alluding to quality of the decision-making process, the testing of different networks and choosing one in particular, which can be done with little discussion (as for the paracetamol-vitamin D case) but also with many committee members contributing (as for the eculizumab case), and stating a preference for the latter form.

DISCUSSION

Resolving the tension underlying health care coverage decisions is difficult by all standards, as visualised by the publicly controversial decisions that are overturned by a direct appeal to the Minister of Health or other public authority. Studying less publicly controversial decisions has been helpful as it illustrates how this identified tension may be resolved in a manner that yields a relatively robust decision. However, it is important to acknowledge that public controversy is 'made' just as much as robustness is (Abelson & Collins, 2009). A lack of public controversy is by no means only dependent on the reasons given for the decision. Many other factors may play a role, a. o. the activities of interest groups, what else occupies the public agenda, how 'publicity-friendly' the case is (McIver & Ham, 2000). Controversy does not just happen; it requires actors to act for controversy to spring into being. Moreover, all decisions are inherently controversial. This, we suggest, is because first, they all manage the perceived tension between the individual and the collective good; between the patient who would like (continued) access to the health care technology and the crowd who would prefer their monthly insurance fee as low as possible (McIver & Ham, 2000; Prainsack, 2018). The second reason all decisions are inherently controversial is that many different elements, many different links, and consequently many different networks are always possible. As a corollary, the model this chapter proposes cannot serve as a prescription for achieving robust, publicly uncontroversial decisions. What it does aim to do, is take a theoretical angle to describe how robustness may be achieved, and as such, it does have potential consequences for possible steps towards more robust decisions. These may include more careful preparatory work both in terms of contents (identification of potential elements from a wide variety of sources) and in terms of the set-up of the hybrid forum setting, enabling everyone to contribute in a way that is sensible to the power differences that are inherent to such decision-making processes. An important resulting recommendation is for further research efforts concerning the relationship, or absence thereof, of network strength and public controversy.

This model is built on previous work on public controversies in STS (Callon et al., 2009; Moreira, 2011; Rip, 1985, 1986) stemming from the premise that in public controversies arguments are explicated that otherwise would remain implicit. This model shows the productivity of tracing

robustness in situations of this inherent perceived controversy, rather than public controversy, and highlights that explicating arguments does indeed also happen in such situations. Though we are by no means the first to focus on the day-to-day, it is a relatively under-explored area in STS' controversy studies (see for an exception, Bal, 1999). We also note the productivity of careful tracing of different elements, and different types of elements, as a way of studying contextualisation processes. This aligns with a recent call to explore decisions of value in health care and how their "inner and outer context" is taken along in these decisions (Williams et al., 2018). However, as the word 'tracing' has been purposely chosen to indicate, it is important to avoid the trap of using context as 'non-explanatory resources' (Asdal, 2012; Asdal & Moser, 2012). By this, Asdal and Moser mean that often, context is said but not *shown* to have affected decisions, and she urges research to precisely indicate how context might have influenced the final decision outcome. In this chapter this has been done through focusing on verbalised, that is, explicated, elements that are added to the decision network. The elements that are (largely) decision-specific, as most of them are, we consider context-derived.

As for the designing and testing of networks, the "plural solutions" to the controversy are tested, according to Callon, by specialists and laypersons together in these hybrid forums, to create 'hybrid knowledge' (Moreira, 2011) or even 'socially robust knowledge' (Nowotny, 2003; Rip, 2010). From these two cases, it becomes clear that the process, for the moment, may not conform to this image completely. What remains to be established empirically, is how much do these practices resemble the hybrid forum ideal already, and how could we make it more so? One important aspect that has been the subject of extensive discussion is opening the decision setting up further to laypersons in addition to the patients already present at (some) appraisal committee meetings (e.g., Abrishami et al., 2017; Baltussen et al., 2016). This is, however, not without its problems. An oft-heard objection to such an appraisal is that it would imply a "tendency to lead to a 'levelling of the epistemological playing field' and to a collapse of the concept of expertise" (Collins et al., 2010; Epstein, 2011; Fischer, 2011). A second objection is the fact that the engagement of the public may be no more than a legitimation exercise (Munk et al., 2016). The third danger, on the other end of the spectrum, is that of deliberations being hijacked by personal interests (Calnan et al., 2017). In response, we would follow Rip and Callon et al. in considering personal interests un-extractable and even constructive to the decision process (Callon et al., 2009; Rip, 1986). The risk of expertise collapsing is somewhat less prominent here because in both decisions examined, 'science proper' was represented by the Institute employee, who had a distinct role in the proceedings, demarcating their expertise also in non-verbal ways. Our data also highlight the appraisal committee's expertise (Calnan et al., 2017; Hughes & Light, 2002; Rooshenas et al., 2015) as being crucial in linking elements; a form of expertise that is not directly based on their respective areas of training and employment. Precisely then because of the distinct role these two types of expert play, the risk of public engagement turning into a legitimation exercise is less easily dispelled than the other two identified problems. In our data, however, a true establishing of networks did indeed happen – it was not just 'for show' in the sense that the decision was already made beforehand. This underlines the potential of encouraging

engagement for the right reasons and in the right ways. One form this may take is through further steps to institutionalise and facilitate appeals procedures, a controversy-sensible step which may also help to counteract these risks.

From publicly controversial decisions that are overturned by a direct appeal to the Minister of Health or other public authority it is clear we need a complex, nuanced picture of how robust coverage decisions are constructed. The picture we have painted of everyday, publicly uncontroversial but inherently controversial decisions values the knowledge that certified experts may bring in the selection and linkage of elements of the decision. It also highlights that other networks, that is, other potentially robust decisions, are available for every decision made. Though no guarantee for publicly uncontroversial decisions, this model shows how coverage decisions may use the productivity of inherent controversy through encouraging the contribution of alternative elements and networks in their decision outcomes. It provides a strong rationale for careful organisation of the deliberative setting and appeals procedure in ways that are easily accessible. It highlights how future research could fruitfully explore the relationship between decision network strength and public controversy.

REFERENCES

Abelson, J., & Collins, P. A. (2009). Media Hyping and the "Herceptin Access Story": An Analysis of Canadian and Uk Newspaper Coverage. *Healthcare policy, 4*(3), e113.

Abrishami, P., Boer, A., & Horstman, K. (2017). 4 Value in Co-Creation: Subjecting Innovative in-Hospital Technologies to Multi-Stakeholder Appraisal. *Public Value of Medical Innovations, 1*, 97.

Aggarwal, A., Ginsburg, O., & Fojo, T. (2014). Cancer Economics, Policy and Politics: What Informs the Debate? Perspectives from the Eu, Canada and Us. *Journal of Cancer Policy, 2*(1), 1-11. doi: https://doi.org/10.1016/j.jcpo.2014.02.002

Asdal, K. (2012). Contexts in Action—and the Future of the Past in Sts. *Science, Technology, & Human Values, 37*(4), 379-403.

Asdal, K., & Moser, I. (2012). Experiments in Context and Contexting. *Science, Technology, & Human Values, 37*(4), 291-306.

Bærøe, K., & Baltussen, R. (2014). Legitimate Healthcare Limit Setting in a Real-World Setting: Integrating Accountability for Reasonableness and Multi-Criteria Decision Analysis. *Public Health Ethics, 7*(2), 98-111. doi: 10.1093/phe/phu006

Bal, R. A. (1999). Grenzenwerk; over Het Organiseren Van Normstelling Voor De Arbeidsplek.

Baltussen, R., Jansen, M. P., Mikkelsen, E., Tromp, N., Hontelez, J., Bijlmakers, L., & Van Der Wilt, G. J. (2016). Priority Setting for Universal Health Coverage: We Need Evidence-Informed Deliberative Processes, Not Just More Evidence on Cost-Effectiveness. *International Journal of Health Policy and Management, 5*(11), 615-618. doi: 10.15171/ijhpm.2016.83

Berkhout, K. (2017). Slikken of Stikken? Het Kan Ook Anders, *NRC Handelsblad*.

Burls, A., Austin, D., & Moore, D. (2005). Commissioning for Rare Diseases: View from the Frontline. *Bmj, 331*(7523), 1019-1021. doi: 10.1136/bmj.331.7523.1019

Callon, M. (1999). The Role of Lay People in the Production and Dissemination of Scientific Knowledge. *Science, Technology and Society, 4*(1), 81-94.

Callon, M., Lascoumes, P., & Barthe, Y. (2009). Acting in an Uncertain World: An Essay on Technical Democracy (G. Burchell, Trans.): Cambridge, MA: MIT Press.

Calnan, M., Hashem, F., & Brown, P. (2017). Still Elegantly Muddling Through? Nice and Uncertainty in Decision Making About the Rationing of Expensive Medicines in England. *International Journal of Health Services, 47*(3), 571-594.

Carlsen, B., & Norheim, O. F. (2005). "Saying No Is No Easy Matter" a Qualitative Study of Competing Concerns in Rationing Decisions in General Practice. *BMC health services research, 5*(1), 70. doi: 10.1186/1472-6963-5-70

Cerri, K. H., Knapp, M., & Fernandez, J.-L. (2014). Public Funding of Pharmaceuticals in the Netherlands: Investigating the Effect of Evidence, Process and Context on Cvz Decision-Making. *The European Journal of Health Economics, 15*(7), 681-695.

Clarke, J. T., Amato, D., & Deber, R. B. (2001). Managing Public Payment for High-Cost, High-Benefit Treatment: Enzyme Replacement Therapy for Gaucher's Disease in Ontario. *Cmaj, 165*(5), 595-596.

Collins, H., Weinel, M., & Evans, R. (2010). The Politics and Policy of the Third Wave: New Technologies and Society. *Critical policy studies, 4*(2), 185-201.

Creswell, J. W., & Poth, C. N. (2017). *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. USA: Sage publications.

Dussauge, I., Helgesson, C.-F., & Lee, F. (2015). *Value Practices in the Life Sciences and Medicine*: Oxford University Press, USA.

Epstein, S. (2011). Misguided Boundary Work in Studies of Expertise: Time to Return to the Evidence. *Critical policy studies, 5*(3), 323-328.

Fischer, F. (2011). The 'Policy Turn'in the Third Wave: Return to the Fact–Value Dichotomy? *Critical policy studies, 5*(3), 311-316.

Franken, M., Le Polain, M., Cleemput, I., & Koopmanschap, M. (2012). Similarities and Differences between Five European Drug Reimbursement Systems. *International Journal of Technology*

Assessment in Health Care, 28(4), 349-357. doi: 10.1017/S0266462312000530

Frey, J. H., & Fontana, A. (1991). The Group Interview in Social Research. *The Social Science Journal, 28*(2), 175-187.

Guindo, L. A., Wagner, M., Baltussen, R., Rindress, D., Van Til, J., Kind, P., & Goetghebeur, M. M. (2012). From Efficacy to Equity: Literature Review of Decision Criteria for Resource Allocation and Healthcare Decisionmaking. *Cost effectiveness and resource allocation, 10*(1), 9. doi: 10.1186/1478-7547-10-9

Hashem, F., Calnan, M. W., & Brown, P. R. (2018). Decision Making in Nice Single Technological Appraisals: How Does Nice Incorporate Patient Perspectives? *Health Expectations, 21*(1), 128-137. doi: doi:10.1111/hex.12594

Hughes, D., & Light, D. (2002). *Rationing: Constructed Realities and Professional Practices*. Oxford, UK: Wiley-Blackwell.

Jasanoff, S. (2003). Breaking the Waves in Science Studies: Comment on Hm Collins and Robert Evans,the Third Wave of Science Studies'. *Social Studies of Science, 33*(3), 389-400.

Jasanoff, S. (2012). Genealogies of Sts. *Social Studies of Science, 42*(3), 435-441.

King, N., Horrocks, C., & Brooks, J. (2018). *Interviews in Qualitative Research*. UK: SAGE Publications Limited.

Kleinhout-Vliek, T., De Bont, A., & Boer, B. (2017). The Bare Necessities? A Realist Review of Necessity Argumentations Used in Health Care Coverage Decisions. *Health Policy, 121*(7), 731-744. doi: https://doi.org/10.1016/j.healthpol.2017.04.011

Kleinhout-Vliek, T., de Bont, A., Boysen, M., Perleth, M., van der Veen, R., Zwaap, J., Boer, B. (2020). Around the Tables – Contextual Factors in Healthcare Coverage Decisions Across Western Europe. International Journal of Health Policy and Management, 9(9), 390-402. doi: 10.15171/ijhpm.2019.145

Lamont, M. (2012). Toward a Comparative Sociology of Valuation and Evaluation. *Annual Review of Sociology, 38*(1), 201-221. doi: 10.1146/annurev-soc-070308-120022

Lehoux, P. (2014). *The Problem of Health Technology*. USA: Routledge.

Lehoux, P., Denis, J. L., Rock, M., Hivon, M., & Tailliez, S. (2010). How Medical Specialists Appraise Three Controversial Health Innovations: Scientific, Clinical and Social Arguments. *Sociology of health & illness, 32*(1), 123-139.

Mackenzie, R., Chapman, S., Salkeld, G., & Holding, S. (2008). Media Influence on Herceptin Subsidization in Australia: Application of the Rule of Rescue? *Journal of the royal society of medicine, 101*(6), 305-312.

Makady, A., Ten Ham, R., De Boer, A., Hillege, H., Klungel, O., & Goettsch, W. (2017). Policies for Use of Real-World Data in Health Technology Assessment (Hta): A Comparative Study of Six Hta Agencies. *Value in Health, 20*(4), 520-532.

Martin, B., & Richards, E. (1995). Scientific Knowledge, Controversy, and Public Decision-Making. *Handbook of science and technology studies, 506*, 26.

Martin, D., Pater, J., & Singer, P. (2001). Priority-Setting Decisions for New Cancer Drugs: A Qualitative Case Study. *The Lancet, 358*(9294), 1676-1681. doi: https://doi.org/10.1016/S0140-6736(01)06714-9

Mciver, S., & Ham, C. (2000). Five Cases, Four Actors and a Moral: Lessons from Studies of Contested Treatment Decisions. *Health Expectations, 3*(2), 114-124. doi: 10.1046/j.1369-6513.2000.00089.x

Moes, F., Houwaart, E., Delnoij, D., & Horstman, K. (2017). Contested Evidence: A Dutch Reimbursement Decision Taken to Court. *Health Economics, Policy and Law, 12*(3), 325-344.

Moreira, T. (2011). Health Care Rationing in an Age of Uncertainty: A Conceptual Model. *Social Science & Medicine, 72*(8), 1333-1341.

Moreira, T. (2012). Health Care Standards and the Politics of Singularities:Shifting in and out of Context. *Science, Technology, & Human Values, 37*(4), 307-331. doi: 10.1177/0162243911414921

Moreira, T. (2013). *The Transformation of Contemporary Health Care: The Market, the Laboratory, and the Forum*. USA: Routledge.

Munk, A. K., Meunier, A., & Venturini, T. (2016). *Data Sprints: A Collaborative Format in Digital Con-*

troversy Mapping: Princeton University Press Princeton, NJ.

Nowotny, H. (2003). Democratising Expertise and Socially Robust Knowledge. *Science and public policy, 30*(3), 151-156.

Nowotny, H., Scott, P. B., & Gibbons, M. T. (2013). *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty*. Cambridge, UK: John Wiley & Sons.

Patera, N., & Wild, C. (2014). Assessment – Appraisal – Decision. Lbi-Hta Decision Support Document Nr.: 72. Wien: Ludwig Boltzmann Institut für Health Technology Assessment.

Prainsack, B. (2018). The "We" in the "Me":Solidarity and Health Care in the Era of Personalized Medicine. *Science, Technology, & Human Values, 43*(1), 21-44. doi: 10.1177/0162243917736139

Ragin, C. C. (2004). Turning the Tables: How Case-Oriented Research Challenges. *Rethinking social inquiry: Diverse tools, shared standards, 123*.

Rip, A. (1985). Experts in Public Arenas. *Regulating industrial risks: Science, hazards and public protection*, 94-110.

Rip, A. (1986). Controversies as Informal Technology a Ssessment. *Knowledge, 8*(2), 349-371.

Rip, A. (2003). Constructing Expertise: In a Third Wave of Science Studies? *Social Studies of Science, 33*(3), 419-434.

Rip, A. (2010). Social Robustness and the Mode 2 Diagnosis. *Science, Technology & Innovation Studies, 6*(1). doi: http://dx.doi.org/10.17877/DE290R-13458

Rooshenas, L., Owen-Smith, A., Hollingworth, W., Badrinath, P., Beynon, C., & Donovan, J. L. (2015). "I Won't Call It Rationing…": An Ethnographic Study of Healthcare Disinvestment in Theory and Practice. *Social Science & Medicine, 128*(Supplement C), 273-281. doi: https://doi.org/10.1016/j.socscimed.2015.01.020

Russell, J. (2017). *The Rationality of Rationing: A Rhetorical Policy Analysis of Deliberations About Resource Allocation in the Nhs.* University of Oxford, UK.

Russell, J., & Greenhalgh, T. (2014). Being 'Rational'and Being 'Human': How National Health Service Rationing Decisions Are Constructed as Rational by Resource Allocation Panels. *Health:, 18*(5), 441-457.

Singer, P. A., Martin, D., Giacomini Bhatia, M. V., & Purdy, L. (2000). Priority Setting for New Technologies in Medicine: Qualitative Case Study. *Bmj, 321*(7272), 1316-1318. doi: 10.1136/bmj.321.7272.1316

Unknown. (2016). Zware Paracetamol Mogelijk Uit Basispakket: Klap Voor Honderdduizenden Patiënten, *RTL nieuws*.

Van Der Aa, E. (2016). Zware Paracetamol En Vitaminen Wellicht Uit Pakket, *Algemeen Dagblad*.

Vuorenkoski, L., Toiviainen, H., & Hemminki, E. (2008). Decision-Making in Priority Setting for Medicines—a Review of Empirical Studies. *Health Policy, 86*(1), 1-9.

Walley, T. (2007). Health Technology Assessment in England: Assessment and Appraisal. *Medical Journal of Australia, 187*(5), 283.

Williams, I., Brown, H., & Healy, P. (2018). Contextual Factors Influencing Cost and Quality Decisions in Health and Care: A Structured Evidence Review and Narrative Synthesis. *International Journal of Health Policy and Management, 7*(8), 683.

Zorginstituut Nederland. (2013). Pakketbeheer in De Praktijk Deel 3 (pp. 1-106). Diemen.

Zorginstituut Nederland. (2016). Reglement Adviescommissie Pakket. In Z. Nederland (Ed.). https://www.zorginstituutnederland.nl/over-ons/publicaties/besluit/2016/12/05/reglement-adviescommissie-pakket-zorginstituut-nederland (accessed 11 February 2020).

Zorginstituut Nederland. (2017). Pakketadvies in De Praktijk: Wikken En Wegen Voor Een Rechtvaardig Pakket. Diemen.

D| Weaving necessity

> Text comes from the Latin *texere,* meaning to weave, and context derives from *contexere,* meaning to weave together or to weave with. (Janssen, 1985, cited in Asdal, 2012)

> [Y]ou should not see that [appraisal process] as too fabricated, I think, because those are actually quite organic, er, organic elements, that discussion, right? (Institute employee 4, interview)

> We have thought too much in recent years about the science of thinking and not enough about the art. (Jacobs, 2017, p. 16)

HOW TO ESTABLISH NECESSITY COLLECTIVELY?

This dissertation was tasked with the quest for a new, more apt, and more usable operationalisation of the necessity criterion in Dutch health care coverage decisions as the process of refining and tightening the establishment of necessity on a collective level was considered to have lagged compared to other criteria. Though not every country employs a formalised necessity criterion, many do utilise qualitative considerations that concern the necessity questionin their deliberations. This question asks: do we think providing this health care technology is necessary on a collective level; or as phrased more specifically in Dutch policy documents and reports, *do we think a claim on the collective solidarity is justified?*

In the Netherlands, decisions concerning the contents of the basic benefits basket are taken by the Minister of Health based on advice formulated by the Dutch National Health Care Institute (in Dutch: *Zorginstituut Nederland*, ZIN, here: the Institute). Many people input into such advice: professionals accumulating and arranging scientific data on effectiveness, cost-effectiveness, and individual severity of illness, building on the work done by many others who have gathered these data; stakeholders who attend and contribute to different meetings that are part of the decision-making process; committees who deliberate and weigh; professionals who write extensive and summative reports. During this multi-cogged process, four criteria are employed so that the final decision may meet all four. Effectiveness and cost-effectiveness of the health care technology in question are first and second, feasibility third, and necessity is the fourth criterion. Using necessity as a criterion has never been as straightforward as the others, and this lack of perceived clarity led to this dissertation.

Over the course of this project (2015-2020), the Institute appears to have moved its position on the necessity criterion and the use of criteria generally. A presentation by an appraisal committee member on an away day (December 2016, field notes #161202) suggested that necessity should be considered an outcome of the whole process rather than input in the form of a criterion. This was followed, a year later, by the publication of the report *Package advice in Practice* (in Dutch: *Pakketadvies in de Praktijk – Wikken en Wegen voor een Rechtvaardig Pakket*) (Zorginstituut Nederland, 2017). Here, the word 'necessity criterion' has largely disappeared in favour of a combination of

severity of illness and own risk considerations. The emphasis of this report lies on argumentations now specifically serving as input for *wikken en wegen* (a Dutch idiom perhaps best translated as a mixture between "deliberation", "weighing", and "hemming and hawing" – the report is now usually referred to at the Institute as the *Wikken en Wegen* report). The report argues that for the acceptance of package advice, it is important that all relevant arguments are visible, that it is clear how they are weighed together, and it is shown which arguments were decisive in the advice and why. "After all," the report states, "new argumentations may always present themselves" (Zorginstituut Nederland, 2017, p. 5). Without concluding that this process of *wikken en wegen* offers a one-to-one replacement for the necessity criterion as previously operationalised, it is important to note the difference in emphasis with the previous *Package management in Practice* reports (College voor Zorgverzekeringen, 2006, 2009; Zorginstituut Nederland, 2013).

My approach to operationalising necessity was influenced by a discrepancy I encountered relatively early on in my research. This discrepancy existed between, on the one hand, the difficulties and critiques displayed by the many scholarly and policy reports that had considered necessity previously and, on the other hand, the relative ease with which Institute employees pinpointed decisions in which necessity was of particular influence. During the early days of my field work, these employees provided a nearly continuous stream of examples: cosmetic surgery, health care aids like walking aids with wheels, maternity care, paracetamol, and vitamin D tablets, physiotherapy, smoking cessation therapies, Viagra, etc., etc. They even often classified these as 'typical necessity questions'. This sensitised me to the difference between argumentations (as present in decision-making practice) and criteria, which seemed mostly visible as such in the decision documents but were stipulated by policy reports to be used in practice. One of the first main questions this project raised in me was: what *is* a criterion exactly? Moreover, I wondered about the fact that many operationalisations had been substantive in terms of descriptions or even checklists (Commissie-Dunning, 1991; Niëns, 2014), even though the process and place of using these argumentations seemed so important in the policy reports (as outlined in the Introduction to this thesis).

From the very beginning – but encouraged by the notable developments in the *Wikken en Wegen* report – I chose to look at the practices of using necessity as a criterion in the form of argumentations, rather than relying on a primarily theoretical approach. The reason was that there seemed to be most clarity in daily practice. The main research question of this dissertation thus became: *how is the necessity criterion used in practice?* The first step entailed the identification of argumentations pertaining to necessity. I purposely widened this first step, the identification of potential argumentations, to scholarly literatures rather than policy reports only. Doing so, I aimed to retrieve a wider variety of argumentations, not only those considered valid enough to become part of a decision justification or rationale. The following steps entailed observing how these argumentations are used in decision-making practice. I did this through a cross-country comparison of practices and decision documents, hoping to gain insight through mirroring 'our' Dutch practices to others', and observations at the appraisal committee meetings enriched by interviews with appraisal committee

members and Institute employees. For the latter two elements, I chose a case approach to gain insight into patterns of social behaviour specifically (Creswell & Poth, 2017; Ragin, 2004).

This discussion starts with the answer to the main research question: how is the necessity criterion used in practice? I will subsequently highlight how using the necessity criterion achieves a societal weighing, conceptualising this societal weighing as *contextualisation practices* while positioning it in relevant Health Services Research (HSR) and Science and Technology Studies (STS) literatures. I will then provide policy recommendations and a reflection on them before ending with the general limitations of this research and a reflection on my research position and role.

ON USING THE NECESSITY CRITERION IN PRACTICE

The unit or form the necessity criterion takes in practice is *argumentations* in favour of, or against, coverage of a health care technology. Argumentations are the explicated reasons pertaining to coverage, generally given on paper or in discussion by anyone reasoning about the potential coverage status of a technology. These necessity argumentations are first and foremost diverse: they may range from scientific calculations on individual severity of illness to more practical considerations such as whether similar treatments are covered or not; from definitions of (non-)illness and what may, or may not, be considered the normal experience to the wider societal impact of coverage (see Chapter 1). Some necessity argumentations are relatively idiosyncratic (as in, specific to the health care technology under consideration) and others more permanent. An example of the first would be the argument that young people should be enabled to step into adulthood with a full set of teeth as this is important for eating apples and daily communication. A (to insiders highly familiar) example of the latter, more permanent type is that as an insurance-based system, to reimburse relatively cheap medicines through health care insurance is considered more expensive than not to do so, meaning that the rule of thumb of 'cheaper than €100 means no coverage' may be applied. In this variability, necessity differs from criteria such as effectiveness and cost-effectiveness, as these generally work with calculations. These may also be, and generally are, under discussion, but the diversity of argumentations and their largely qualitative nature are the first characteristics that make necessity stand out.

The selection of necessity argumentations differs per health care technology under consideration and is crucial to the use of the necessity criterion in practice. As Chapter 1 shows, the reason is that not all necessity argumentations are considered equally valid for the decision on every health care technology. This makes earlier efforts to operationalise necessity in the form of a checklist (Niëns, 2014) relatively difficult to use, as it indeed turned out to be (interviews appraisal committee member 5, Institute employee 6, March 2015). It also sets necessity again in stark contrast to more classic coverage decision criteria such as effectiveness and cost-effectiveness (cf. Stolk & Poley, 2005). Others have noted, over fifteen years ago, that the use of what they termed considerations pertaining to necessity was inconsistent, which made it a "problematic notion" (Hoedemaekers

& Oortwijn, 2003). This dissertation fully supports the conclusion of inconsistency in terms of necessity argumentation use, though expressly not the valuation thereof.

When comparing the use of necessity argumentations across countries, it becomes clear that not every country has formalised necessity in a criterion to the same degree. This does not mean, however, that necessity argumentations are absent elsewhere. In fact, the use of necessity argumentations is widespread, at least in the Western world but also not absent from other middle- and high-income countries such as Argentina, Israel, Japan, and Korea (Chapter 1). The reason it may perhaps appear a quintessentially Dutch notion is, at least in part, the strong tradition of explicating and operationalising these argumentations, as evidenced by the policy reports described in the Introduction but also by scholarly work in this area (Hoedemaekers & Oortwijn, 2003; Stolk et al., 2002). Chapter 2 delves in further to show how in Belgium, England, and Germany, necessity argumentations are both recognised and used by decision makers. The Netherlands has both a comparatively high level of formalisation of necessity and a broad set of necessity argumentation usage. This holds generally across the decisions on the four health care technologies studied but is especially visible in the individual cost consideration. This consideration indicates a lower necessity of coverage of health care technologies with a low price and is unique to the Netherlands in this dataset. The high level of explicated and formalised necessity argumentations confirms Dutch coverage practice as an especially fruitful research site.

From the cross-country comparison, it also becomes clear that the combination or *pattern* of argumentation use is of high importance, as similar patterns of argumentation use lead to similar health care coverage decisions in the countries studied. In Chapter 2, I highlight similarities between argumentation patterns in specific decisions, such as the English and Dutch decisions on nivolumab, benzodiazepines, and smoking cessation therapies. Chapter 1 resonates this conclusion as it not only demonstrates that some argumentations are used primarily in favour of, and others primarily against, coverage but also indicates similarities between the argumentation patterns for cancer drugs and orphan drugs, which are very different from the argumentation patterns for cosmetic surgery, Viagra, infertility treatments, obesity treatments, and smoking cessation therapies. This research thus underlines the previously-noted but not well-studied importance of clustering or patterning of argumentations (Lehoux et al., 2010; Martin et al., 2001; Singer et al., 2000). Notably, these argumentation patterns could well coincide with potential different decision networks with varying robustness (Chapter 4). This indicates that exchange between countries on more qualitative aspects of health care coverage decisions, in addition to ongoing efforts on quantitative aspects, may well be a fruitful endeavour.

Necessity argumentations are used 'around the table', that is, in a *deliberative* setting, in all four countries studied in Chapter 2. In the Netherlands, this primarily takes the shape of the appraisal phase of coverage decisions. The appraisal phase generally follows the assessment phase, in which the relevant scientific input is determined, but it may or may not be a distinct moment in time and space (Oliver et al., 2004; Patera & Wild, 2014; Walley, 2007). The goal of the Dutch appraisal, which is indeed separated from the assessment, is to provide a *societal weighing* of this input, similar to the

former social value judgements in England (Culyer & Rawlins, 2004; NICE, 2008; Shah et al., 2013; Zorginstituut Nederland, 2017). This appraisal is achieved by the appraisal committee (in Dutch: *Adviescommissie Pakket*, ACP), a committee of eight to ten experts with professional backgrounds such as pharmaco-economics, health care ethics, and patient sciences (art. 14, Zorginstituut Nederland, 2016).

To describe how necessity is constructed and what constitutes societal weighing expertise therein, I the appraisal committee meeting in detail in Chapters 3 and 4. From this, I conclude that societal weighing expertise has two primary facets in this dataset. First, it takes the form of receiving input and allowing this input (in the form of explicit criteria presented by Institute employees and contributions by patient (representative)s) to steer and challenge the process to increase the quality of the decision. This dynamic of perceived increased quality of decisions due to the presence and contribution of patient (representative)s, I would suggest, has two primary explanations. First, unlike other criteria noted above, necessity is the type of consideration that may more easily also be contributed by laypersons and those with experiential expertise (Chapter 1). In this way, the variety and number of argumentation types may be increased and the resulting decision potentially strengthened (Chapter 4). The committee members themselves expressed the second explanation, namely how, primarily for expensive medicines with a relatively unfavourable cost-effectiveness ratio, the committee was forced to consider matters extra carefully and formulate their negative coverage decision advice especially diligently, when faced with patients and/or their representatives (Chapter 3).

The second facet of societal weighing expertise is the deliberation and the formulation of the advised decision and recommendations. This deliberation is where the pattern of necessity argumentations that is to become part of the final coverage decision is established. The appraisal committee does so by bringing in new argumentations, derived from many sources ranging from scientific reports to newspapers and beyond, and *weaving the argumentations together*. This weaving serves as a metaphor for the verbal combining of different case-specific and actively-integrated argumentations considered pertinent to the case at hand (Chapter 2). Different argumentations may be woven together in different ways, of which I conceptualise three in Chapter 4. First is the articulation of links between argumentations, whereby different argumentations are brought together as both relevant to a positive or negative decision. Examples include linking the coverage of maternity care not to preparing *beschuit met muisjes* but to detecting risky situations. Using black-boxed links is the second. Black-boxed links are previously-formulated connections, of which the negative decision based on an unfavourable cost-effectiveness ratio appears the most common, described by some as "simply stamping the file" (appraisal committee member 6, personal communications). The third and final way of combining argumentations is broadening the scope of networks, i.e., including previously unconnected matters to a decision under construction. This includes comparing it to a previous decision, in the sense that 'entering adulthood with good front teeth' should be covered just as 'entering adulthood without an unwanted pregnancy' is. Broadening the scope of networks also comprises the formulation of specific recommendations. These recommendations

may be aimed at actors as diverse as the Minister of Health and treatment expertise centres abroad. The intended end-product, an advised decision plus recommendations, I conceptualise as a robust decision, meaning able to withstand pressure 'out there'.

In sum, Chapters 1 and 2 describe how the necessity question – is a claim on the collective solidarity justified? – is answered in practice by using argumentations. These argumentations are numerous and diverse, and their perceived validity differs per decision. In some countries, like the Netherlands, these argumentations are formalised to a high degree into a criterion, whereas in other countries these argumentations are used without such a high level of formalisation. In the Netherlands, a high level of formalisation does coincide with a high diversity of argumentation types. Moreover, necessity argumentations are used in patterns, with similar patterns leading to similar decisions across countries. Finally, which argumentation types to use is determined in deliberation, 'around the table', which generally occurs in the Netherlands in the appraisal phase with the explicit purpose of achieving a societal weighing. Zooming in on the Dutch appraisal, Chapters 3 and 4 describe what such a societal weighing comprises: allowing explicit criteria and patient (representative)s' contributions to steer and challenge the process, and subsequently, bringing in new argumentations, combining them in different ways, (namely through linking argumentations, using black-boxed links, and broadening the scope of decisions), making the decision and formulating recommendations. This shows how argumentations and other decision elements are combined into a robust decision.

SOCIETAL WEIGHING AS CONTEXTUALISATION PRACTICES

The societal weighing that happens in appraisal, identified under the previous heading as the primary place of use of necessity argumentations, serves to set the decision-making process and outcome in context, to contextualise it. I will show in this section that societal weighing entails first, 'bringing the outside in', and second, placing the decision 'back into society'. I will then set these findings in the wider Health Services Research and Science and Technology Studies literatures relevant to coverage decision-making practice, explicating my contribution by describing societal weighing as *contextualisation practices*.

Bringing the outside in

In the process of societal weighing, 'the outside' is actively brought in. This happens in two primary ways, first, through the patients, patient representatives, and/or other stakeholders that are present in the room (or have been present in the scoping session), and second, through argumentations brought in by the appraisal committee members. I will discuss these in turn.

First, the patient, patient representatives, and/or other stakeholders strongly impact the processes of societal weighing as they represent the outside. They arrive in the scoping or appraisal committee meeting literally from the outside: they use the visitors' entrance, sit in the audience section (see

Chapter 3), and are here for this decision only. In this dataset on appraisals, these groups were almost exclusively visible in the eculizumab case, where a mother described the roller coaster their family had experienced since their daughter's diagnosis with aHUS. However, the effect of patient (representative)s on the deliberations were a primary topic in the interviews with the committee members. Their contributions are considered to steer or guide the deliberative process in terms of challenging the committee to substantiate their decision well. Moreover, other stakeholders contributed in several cases at other moments, such as pharmacists in the paracetamol-vitamin D case drawing attention to vulnerable groups, and the association of dentists who set the front teeth replacement therapy decision on the agenda and inputted later on as well.

The second way in which the outside is brought in is through the committee members themselves, who derive argumentations from a wide variety of 'outside sources'. These sources include but are not limited to newspapers (as in the case of maternity care, where reference was made to *beschuit met muisjes smeren*) and previous decisions (as in the cases of front teeth replacement therapy, where it was linked to the coverage status of contraceptives for under-18s, and eculizumab, with the earlier decision for PNH patients). Naturally, argumentations may also come from the contributions noted above, as the pharmacists who contributed during the scoping session for paracetamol-vitamin D tablets, whose argumentations were repeated by an appraisal committee member.vv The committee, then, may add an argumentation from the outside to a (collective) decision network and in this way ascertain and, if successful, ensure its validity in this case.

In societal weighing, the outside is actively brought in through the presence and contributions of patients, their representatives, and/or other stakeholders, which steer or guide the deliberations, and through the committee members themselves as they bring in new argumentations from outside sources and add them to decisions.

Placing the decision back outside

Placing the decision back outside is likewise achieved through two interlinked ways. The first element is that the committee seeks to make a *robust* decision, conceptualised by Rip as being able to withstand pressure in particular outside settings, achieved through careful bringing together of argumentations into a solid justification or rationale. The second element is the addition of recommendations.

The first element, making a robust decision, entails for the committee to bring argumentations and other decision elements together with the societal context in mind, which means here to make reasonings explicit. Specifically, the explication of argumentation types for reasons of providing a justification or rationale is important here. Such explication into networks of argumentation types cannot be performed randomly, seen first in the fact that many argumentation types appear to be used for either positive or negative coverage decisions. They tend to come in patterns, as shown by the similarities between several argumentation patterns in the cross-country comparison, such as the English and Dutch decisions on nivolumab and smoking cessation therapies. It is also shown

by the ways alternative decision networks are explicated, such as the appraisal committee member who professed to be able understand that *beschuit met muisjes smeren* should not fall under the collective solidarity, as well as the committee member who posited that safeguarding the disposable income of vulnerable groups is not the prime responsibility of the basic benefits basket. Both were done to benefit the reason-giving: to provide a justification or rationale displaying that these matters had been thought about, though dismissed.

The second way in which the committee places the decision back outside is through recommendations to other stakeholders. These recommendations comprise, in this dataset, continued work on indication criteria (e.g., the maternity care case, but also the English and Dutch decisions on benzodiazepines and the German decision on walking aids with wheels), price negotiations (nivolumab case), but also cooperation with other research centres (eculizumab case) and for the Minister to "to think carefully about the prescription rule" (paracetamol-vitamin D case). Importantly, giving recommendations is not part of the formal remit of this committee; they actively branch out when they give such recommendations. These practices also serve to make the decision more robust, even specifying the outside settings in which the decision is to have an effect.

> *Societal weighing thus secondly entails placing the decision back outside through making careful justifications or rationales for decisions and making them explicit to benefit those outside, and through formulating recommendations for specific outside settings.*

Contextualisation practices

In this section, I would like to concretise how conceptualising societal weighing as contextualisation practices builds on, and contributes to, the HSR and STS literatures.

First, as outlined in the Introduction, scholars in HSR have described committees like the Dutch appraisal committee as taking decisions flexibly, humanely, with sensitivity towards emotions and preferences of the recipient(s) of the health technology under consideration as well as those making the decision (Hughes & Light, 2002; Mechanic, 1997; Russell & Greenhalgh, 2014). This is combined with adhering to decision criteria and procedures. These committees are especially good at doing both (Hughes & Doheny, 2011; Jenkings & Barber, 2004; Russell, 2017). Russell has designated this combination as experts displaying *pragmatic rationality* (Russell, 2017, following a.o. Aristotle; Russell & Greenhalgh, 2014).

HSR's conceptualisation of pragmatic rationality as human (responsive to patients and others) and rational (responsive to criteria and procedures), this dissertation shows, has a distinct outworking when it comes to societal weighing. In societal weighing, these patients, their representatives, and/or stakeholders are representing the societal *context* in the deliberative setting. During the deliberations, the appraisal committee feels they need to keep their distance (cf. Moreira, 2012) but that the presence of patient (representative)s does give them "handles" for substantiating their position. In this sense, this outside that is brought in by patient (representative)s being present, this context, follows STS scholar Asdal as it "conditions or enables a specific utterance to happen"

(Asdal, 2012, p. 388). As such, it is primarily the people that are present that have this type of impact (cf. Wallenburg et al., 2019), and this is perceived to increase the quality of the decision outcome. Notably, this links the presence of the patients, their representatives, and/or other stakeholders, who represent the outside, directly to the quality of the decision justification or rationale, which is for the benefit of the outside. It is a double form of contextualisation. Deliberations, decisions, and justifications or rationales being shaped by these actors in this way I thus denote as contextualisation practices.

Second, STS has classically described pragmatic rationality as taking into account circumstances relevant to that unique situation in order to achieve robust outcomes in particular settings (Rip, 1985, 1992). Rip specifies that pragmatic rationality is crucial in the absence of scientific and socio-political consensus, in line with more recent studies that highlight such rationality in situations of high uncertainty (Calnan et al., 2017; Moreira, 2011). Within STS, one-off controversy has been deemed an important area of study. The reasoning is that many implicit argumentations become explicated in confrontational settings especially (Callon et al., 2009; Moreira, 2011, 2013).

As noted in Chapter 4, this study shows that such explication also happens, at least in part, in less controversial decisions. Moreover, it is not just explication of argumentations; the committee carefully brings them together by bringing the outside in (deriving from newspapers and the like) and adding recommendations to stakeholders, who are in daily life positioned outside. Argumentations to be factored into the decision are fragile during the process and needs to be solidified or linked in some way, which happens in the decision-making situation (Callon et al., 2009; Nowotny, 2003; Rip, 1992). Part of the contextualisation of the decision thus involves deciding which argumentation types are, and which are not, taken along in the final decision: around the table is where the active *integration* of the outside argumentations happens. Moreover, it is where the recommendations are formulated, which is a specification of actively making a decision robust *in certain outside contexts*. Recommendations specify who should do what to 'make this work', to heighten the quality of care. As indicated earlier, previous work has not engaged much with giving recommendations, and this is therefore an important nuance: recommendations make a decision robust in certain contexts. Here we thus also see a double contextualisation movement: active integration of outside argumentations, combined with active formulation of recommendations to benefit outside contexts.

Concluding, conceptualising societal weighing as contextualisation practices focuses our gaze on two elements. First, it highlights the impact of patients, their representatives, and/or other stakeholders as representing the outside in terms of the quality of the decision justification, which is for the benefit of those outside; second, it shows the expertise of the committee in terms of choosing the right outside argumentations, combining them, and formulating recommendations, all deriving from or aimed at this same outside, i.e. the societal contexts as conceived of for that specific health care technology and coverage decision.

Societal weighing is achieved through contextualisation practices, which specifically conceive of patients, their representatives, and/or other stakeholders as representing outside contexts which

condition or enable utterances to happen and thereby increasing the quality of the decision justification or rationale for the benefit of those outside. Second, they comprise actively integrating argumentations deriving from outside contexts, showing that explication also happens in less controversial decisions, and adding recommendations to stakeholders which are positioned in certain outside contexts.

EXPERTS AS GUARANTORS OF DECISION QUALITY

In this section, I will briefly explore the role of the committee in terms of experts as guarantors of decision quality in view of the different 'types' of societal weighing that this dissertation has described, and how this differs from the traditional types of guarantors described in the Introduction, namely through adherence to procedural or substantive criteria.

Though all societal weighing processes, characterised by contextualisation practices, display elements of bringing the outside in and placing the decision back outside, the way they do so varies significantly. Societal weighing expressly does not look the same for every decision. The societal weighing takes a different shape every time and this is most obvious in how the committee assumes different roles and by extension in how the collective solidarity through the basic benefits basket assumes different shapes. Examples that stand in stark contrast include: "we want to make something possible here" (eculizumab case) or "simply stamping the file", meaning that the committee functions as an enabler of price negotiations (Dutch and English nivolumab decisions). Another apparently contradictory set would be 'not responsible for the disposable income of certain groups' (paracetamol-vitamin D case, as there are other responsible organisations, but also the Dutch walking aids with wheels decision) contrasted with 'responsible for children and young people' (maternity care and front teeth replacement therapy cases). Many may conceive of these differences as problematic. I will proceed to argue the opposite, namely that choosing the right role and thus demarcating the extent of the collective solidarity is exactly what the appraisal committee is supposed to do.

The robustness of a decision is dependent on the setting(s) in which the decision is to play a role, and the particulars of these settings are different every time. The societal context in this dissertation is expressly not used as a non-specific explanatory resource (Asdal, 2012; Asdal & Moser, 2012): *the societal context varies per decision as it is differently conceived of for that specific health care technology and coverage decision*. Let me demonstrate this. The examples above showcase different patients of differing ages, differing clinical pictures, but also differing ways in which the care is provided, different stakeholders and divisions of responsibilities, different initiatives and possibilities in terms of research but also in terms of price negotiations. Naturally, the effectiveness and cost-effectiveness of a health care technology differ every time as well but the role these criteria may play and their perceived validity do, in principle, not vary to the same extent across decisions. The committee, then, does well to be aware of the particulars of a decision's specific societal contexts and to take

them along actively in the decision-making process. In this active taking-along, this integration, it is important to note that the people present in the appraisal committee meeting, the patient (representative)s, other stakeholders, Institute employees, and the committee members themselves, all contribute to the fact that *not anything goes*. Instead, delineating the collective solidarity in the right way for a particular health care technology requires expertise in terms of societal weighing, evidenced by this specific set of contextualisation practices: bringing the outside in and placing the decision back outside in a careful, appropriate, and well-legitimated manner.

> *Not all societal weighing processes look the same: in fact, they vary significantly. This is shown by the fact that the committee takes different roles in the deliberations, and that correspondingly, the extent of the collective solidarity is demarcated differently. This, I argue, is not a problem – rather, it is exactly how it should be. Not only varies the health care technology per decision but also the contexts that are perceived as relevant are significantly different each time. This means that in expert contextualisation practices, through which a high-quality societal weighing is achieved, these different contexts are taken into account well. This results in carefully made, appropriate, well-legitimated, robust decisions.*

WEAVING NECESSITY

Above, I have described contextualisation practices as the processes of using argumentations to achieve an expert societal weighing. I have chosen the metaphor of weaving to denote these contextualisation practices and I will spend some words on this metaphor here. The contextualisation practices described are evidently more than muddling through, even elegantly; they require expertise to achieve well. They achieve more than rationality, even pragmatic rationality; they achieve a decision that is considered well-rounded and well-grounded. The metaphor of weaving, evoking notions of a structured, organised craft, with differing substrates in terms of colours and thickness but clear boundaries in terms of what may be achieved, I trust, helps bring out the art of making health care coverage decisions. Previous operationalisations of the necessity criterion, whether in the form of checklists or broadly-defined considerations, seem to have missed this aspect. I would argue they have thought more about the 'science' of decision making, and less about the art (Jacobs, 2017).

The use of the necessity criterion in practice, I would like to characterise as 'weaving necessity'. Necessity argumentations, which are many, case-specific, actively integrated during deliberations around the table, and varied in level of formalisation but used in patterns, ensure that decisions are not made without reference to relevant societal contexts. Weaving necessity, then, entails using these carefully-selected argumentations in a way that achieves a societal weighing (in terms of both bringing the outside in and placing the decision back outside), delineating the collective solidarity

in a way that is well-legitimated and sensitive to the case at hand, in a way that yields a decision that is robust in contexts that are considered relevant.

> *Weaving necessity is a metaphor I use to evoke the art of decision making, showcasing that necessity argumentations, which are many, case-specific, and patterned, are actively integrated into decisions through these contextualisation practices. Using necessity argumentations thus aids achieving a high-quality societal weighing through delineating the collective solidarity in an appropriate, robust manner.*

POLICY RECOMMENDATIONS

I would like to provide three specific policy recommendations and a brief reflection on all three. Before I begin, it is clear that policy practice is also on the move while research happens (Zorginstituut Nederland, 2017), and it may well also point to how the two worlds intertwine. This may happen through informal conversations and more formal presentations; the recent *Wikken en Wegen* report is written by my primary point of contact at the Institute, who also co-authored Chapter 2. Three specific elements of this report as describe above resonate strongly with this dissertation. First, the renewed interest for reference to the societal context. Second, the non-exhaustiveness of a list of criteria: additional relevant argumentations are always possible. Three, the importance of a strong justification or rationale to back up the advised decision. In line with the above, I would advise the Institute to continue on this path by following three recommendations.

Contextualisation practices entail first, bringing the outside in, in which patients, their representatives, and/or other stakeholders play a vital role. Consequently, my first recommendation is to take steps to *invite more different perspectives* into the deliberative process in the appraisal. This tallies with Moes' reasoning that this values especially patients in their capacity as knower (Moes, 2019) and the scoping sessions currently being institutionalised at the Institute (Zorginstituut Nederland, 2017). In my words, this entails bringing more different and more different types of argumentations into the decision-making process. To this end, I would encourage reflection on how to achieve that and whom to involve, as well as exchanges between HTA agencies on more qualitative aspects of health care coverage decisions. I would add that having more actors present in the decision-making setting might also contribute to the realisation of the rest of the decision through linking argumentations together into decision networks. This would also thus potentially facilitate a more rigorous testing step, as the appraisal committee currently tests these decision networks alone. This would enable research on how different forms of expertise and experiential knowledge work differently in terms of 'opening up the decision network (a suggestion which I owe to Professor Tiago Moreira). As a corollary, I suggest *institutionalising an improved appeals procedure*, as it will open opportunities to contribute to not only ongoing decisions but also to

those taken in the past, perhaps redressing at least part of the power differential inherent to these decisions.

Second, contextualisation practices comprise setting the decision back outside. That includes, as stated above, the involvement of others not only in the scoping phase but also in the testing during the appraisal phase. Moreover, as the justification or rationale is essential herein, I would suggest that the final advised coverage decision would benefit from *including alternative decisions that were considered but not chosen*. I would suggest that if for example both a positive and a negative decision were considered, or two variations on a positive decision, both potential decisions are displayed in the final advised decision including the argumentations that were considered to back up these alternatives. The reason is that this inclusion of alternative decisions would helpfully show that the concluding decision was not consensual, spelling out rather than obscuring differing points of view (Mouffe, 1999, 2011; Rip, 1986). Although a potential controversy may, of course, still follow, a relatively robust decision outcome will benefit from the transparency gained by the inclusion of alternative decisions that were also considered.

Third, I would advocate for the Institute to consider ways to *actively find controversy and invite it in* (again not just in the appraisal phase but more generally), as an additional way to bring the outside in. Callon et al. suggest:

> [Controversies] should be encouraged, stimulated, and organized. There are overflows everywhere. They produce the fabric of our individual and collective lives. (Callon et al., 2009, p. 257)

They see overflows, that is, places of controversy, everywhere. This yields another recommendation for the Institute: actively finding controversy that may be brewing 'in society', and drawing it in. This is another, more proactive form of contextualisation, and would perhaps take the form of horizon scanning not just for expensive medicines but for controversy in the making. Examples might include organising a conversation on the inclusion of menstrual cups in the basic benefits basket as suggested by a recent petition (https://petities.nl/petitions/menstruatiecup-in-de-basisverzekering, accessed 19 November 2019). This actively inviting in controversy in the making also yields a task for researchers: designing and experimenting with methods to identify (hidden) controversy, in addition to existing methods to map them (e.g., Marres, 2015; Munk et al., 2016).

All three recommendations may appear to rely on a somewhat idealistic and rationalistic description of deliberation, and as such, clash with the largely organic process that I describe in terms of coming to a decision through weaving. Regarding the second recommendation for example, in the contextualisation process, one potential outcome takes precedence over another, the committee considers it more worth investing 'weaving time' into, and as such, the decision that is not taken will never be as carefully woven as the decision that is. On the first and third recommendation, this weaving process does by no means guarantee all elements to be taken along by those who weave – and thus far, the experts at contextualisation have been the appraisal committee, which has given the deliberative situation another dimension to the already notable power differential in the room.

Finally, it does not explicitly acknowledge the antagonism that is and will be present, in part due to the inherent tension between the individual and the collective (Moreira, 2011), an antagonism that is not to be obscured by consensus (Mouffe, 1999, 2011). Thus far, the weaving process and robust outcomes of decisions have resonated with and strengthened one another, but here these conceptualisations do clash. I would argue that robustness takes precedence here. This means that even though this alternative decision is less carefully woven, which will probably take the form of being less precise, less extensive, and less well-formulated, it is worth including it in the decision outcome nevertheless, for reasons of transparency. These recommendations, then, ultimately hope to provide concrete steps to lessening the power differential somewhat, risking antagonism for the relatively robust and transparent outcomes it may produce.

STRENGTHS AND LIMITATIONS

This study is the first to produce an overview of argumentations pertaining to necessity. It is strong in terms of showing how these argumentations are used, as it does so through analysing observations/audio recordings, documents, and interviews in the Netherlands but also abroad. It interprets a relatively large number of cases (four Dutch appraisals and sixteen further decision outcomes) and finds commonalities in argumentation use across those. Naturally, because of time constraints, some in-depth familiarisation may have been lacking.

Moreover, my efforts have largely been focused on the appraisal phase and the documentation surrounding it, but contextualisation may be conceived of as much more than just the appraisal. This is visible, for example, in the scoping session and the Institute employees who write up the discussion documents. Many argumentations that were prepared were, however, verbalised, explicated, in the appraisal meeting, making it a prime locus for this study. Within the appraisal meetings, I have focused on explicit rather than implicit meaning-giving. This was, as reflected on below, formed by both my personal research interests as well as the practicalities of using explicated argumentations as sensitising concepts. It does remain likely that in this way, I have put less emphasis on more implicit forms of contextualisation. Similarly, I have had a relatively narrow focus on necessity. How these different contextualisation practices intertwine and what more implicit forms look like would benefit from further research.

REFLECTION ON PERSONAL RESEARCH POSITION AND ROLE

In answering the research questions, it is important to be aware of the epistemological angle I have approached them from and how this may have impacted the data collection and interpretation.

My reflections are on epistemology and specifically on how it impacts doing research in terms of methodology, conceptualisations, and conclusions. With my background in the natural sciences

(biomedical research to be precise), I am used to showing *care* towards my research subject(s). Counterintuitive as this may sound to some, cultured cells require diligent care (cleaning, feeding, and the like), and stem cells, in particular, are fussy research subjects, necessitating the researcher to give up Saturday afternoons to care-giving tasks. Naturally, this imperative of taking care holds even more for those with human research subjects and, according to Latour, especially of those who adopt a constructivist stance, as I have done in this research (Latour, 2004). To specify, by a constructivist stance, I understand that 'facts', scientific or otherwise, are made, knowledge is generated, in the sense that knowledge is *mediated* in some way.

This means first and foremost that my academic work has been mediated: I have been as involved in and as responsible for data generation as the people and practices I studied, and likely even more so. This is visible in the construction of the twenty necessity argumentation types, which are mine alone (but checked with a co-reader). It is no less true for the conclusions of Chapters 2 and 3, even though these were member-checked rigorously. Law and Singleton state that research means selecting, and that by writing up data into chapters, I betray part of the data that were available to me (Law & Singleton, 2013). I would have to agree that other conclusions, or at least other nuances, would indeed have been possible. I have two examples of how my personal preferences have shaped my research. First, as a researcher, I am interested in investigating what I consider tangible: I find it difficult to convincingly show how external factors may implicitly affect decisions (Asdal, 2012; Asdal & Moser, 2012). This has impacted the way I have opted to follow the argumentation types across the different observed meetings and documents (Chapters 2-4), rather than follow interviewees' ideas of how certain contexts, such as the presence of innovative pharmaceutical industry or David Cameron's desire to be re-elected, both of which were mentioned by interviewees, may have had an indirect impact on certain coverage decisions. Second, I have been shaped by a book called 'How People Think' (Jacobs, 2017), where Alan Jacobs convincingly argues (following Daniel Kahneman and Jonathan Haidt, amongst others) that people make decisions intuition-first. This is followed by a rationalisation of this intuition by providing reasons for why this intuition is right. I also believe, however, that this process of providing reasons is important for the legitimacy of public decisions especially. I think that decisions that are made by a direct appeal to a Minister of Health, who does not have to account for his or her decisions in the same way, are less valid – as visible in Chapter 4 especially. Overall, during its coming together and when it is finished, academic work is performative, it *does* something. I am therefore not surprised to see ongoing parallels between my work and the policy developments at the Institute (Law & Singleton, 2013).

On taking care as a constructivist, Latour specifies that:

> The critic is not the one who debunks, but the one who assembles. The critic is not the one who lifts the rugs from under the feet of the naive believers, but the one who offers the participants arenas in which to gather (…) one for whom, if something is constructed, then it means it is fragile and thus in great need of care and caution. (Latour, 2004, p. 246)

From this, I gather that to believe that the research subject is constructed and to show how it is constructed, how it is made, is to expose its vulnerabilities. For me personally, this only truly acquired meaning after a particularly vivid member check, in which I appeared to have misinterpreted the deliberations at the appraisal committee I studied. Sayer gives hands and feet to this idea of taking care when he says in his book 'Why Things Matter to People' that it is possible for social scientists to over-theorise to such an extent that the people studied *do not recognise themselves* (Sayer, 2011). He argues that social scientists should aim to remain close to the interpretation of those studied. These two ideas, of showing care and of remaining close to the interpretation of the research subjects, have helped me find direction for my dissertation. These ideas have impacted the methodology, some of which was chosen in collaboration with the research subjects (Chapter 2). These ideas may also have, perhaps, impacted the theoretical concepts employed to give meaning to the data gathered. After the deconstruction phase of studying the appraisal committee meetings and the advised decisions, I have chosen to conceptualise the *robustness* of health care coverage decisions. I believe this does justice to the everyday work at the Institute, and I believe it may potentially contribute constructively to future efforts.

REFERENCES

Asdal, K. (2012). Contexts in Action—and the Future of the Past in Sts. *Science, Technology, & Human Values, 37*(4), 379-403.

Asdal, K., & Moser, I. (2012). Experiments in Context and Contexting. *Science, Technology, & Human Values, 37*(4), 291-306.

Callon, M., Lascoumes, P., & Barthe, Y. (2009). Acting in an Uncertain World: An Essay on Technical Democracy (G. Burchell, Trans.): Cambridge, MA: MIT Press.

Calnan, M., Hashem, F., & Brown, P. (2017). Still Elegantly Muddling Through? Nice and Uncertainty in Decision Making About the Rationing of Expensive Medicines in England. *International Journal of Health Services, 47*(3), 571-594.

College Voor Zorgverzekeringen. (2006). Pakketbeheer in De Praktijk. Diemen.

College Voor Zorgverzekeringen. (2009). *Pakketbeheer in De Praktijk 2*. Diemen: College voor Zorgverzekeringen.

Commissie-Dunning. (1991). Kiezen En Delen. *Rapport van de Commissie Keuzen in de Zorg*.

Creswell, J. W., & Poth, C. N. (2017). *Qualitative Inquiry and Research Design: Choosing among Five Approaches*. USA: Sage publications.

Culyer, A. J., & Rawlins, M. D. (2004). National Institute for Clinical Excellence and Its Value Judgements. *British Medical Journal (BMJ), 329*, 224-227.

Hoedemaekers, R., & Oortwijn, W. (2003). Problematic Notions in Dutch Health Care Package Decisions. *Health Care Analysis, 11*(4), 287-294. doi: 10.1023/B:HCAN.0000010057.43321.b2

Hughes, D., & Doheny, S. (2011). Deliberating Tarceva: A Case Study of How British Nhs Managers Decide Whether to Purchase a High-Cost Drug in the Shadow of Nice Guidance. *Social Science & Medicine, 73*(10), 1460-1468.

Hughes, D., & Light, D. (2002). *Rationing: Constructed Realities and Professional Practices*. Oxford, UK: Wiley-Blackwell.

Jacobs, A. (2017). *How to Think: A Survival Guide for a World at Odds*: Currency.

Jenkings, K. N., & Barber, N. (2004). What Constitutes Evidence in Hospital New Drug Decision Making? *Social Science & Medicine, 58*(9), 1757-1766.

Latour, B. (2004). Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern. *Critical inquiry, 30*(2), 225-248.

Law, J., & Singleton, V. (2013). Ant and Politics: Working in and on the World. *Qualitative Sociology, 36*(4), 485-502.

Lehoux, P., Denis, J. L., Rock, M., Hivon, M., & Tailliez, S. (2010). How Medical Specialists Appraise Three Controversial Health Innovations: Scientific, Clinical and Social Arguments. *Sociology of health & illness, 32*(1), 123-139.

Marres, N. (2015). Why Map Issues? On Controversy Analysis as a Digital Method. *Science, Technology, & Human Values, 40*(5), 655-686.

Martin, D., Pater, J., & Singer, P. (2001). Priority-Setting Decisions for New Cancer Drugs: A Qualitative Case Study. *The Lancet, 358*(9294), 1676-1681. doi: https://doi.org/10.1016/S0140-6736(01)06714-9

Mechanic, D. (1997). Muddling through Elegantly: Finding the Proper Balance in Rationing. *Health Affairs, 16*(5), 83-92.

Moes, F. (2019). *Experts and Idiots: An Ethnography of Epistemic Practices in Dutch Social Health Insurance.* Maastricht.

Moreira, T. (2011). Health Care Rationing in an Age of Uncertainty: A Conceptual Model. *Social Science & Medicine, 72*(8), 1333-1341.

Moreira, T. (2012). Health Care Standards and the Politics of Singularities:Shifting in and out of Context. *Science, Technology, & Human Values, 37*(4), 307-331. doi: 10.1177/0162243911414921

Moreira, T. (2013). *The Transformation of Contemporary Health Care: The Market, the Laboratory, and the Forum*. USA: Routledge.

Mouffe, C. (1999). Deliberative Democracy or Agonistic Pluralism? *Social research*, 745-758.

Mouffe, C. (2011). *On the Political*: Routledge.

Munk, A. K., Meunier, A., & Venturini, T. (2016). *Data Sprints: A Collaborative Format in Digital Con-*

troversy Mapping: Princeton University Press Princeton, NJ.

Nice. (2008). *Social Value Judgements - Principles for the Development of Nice Guidance*. https://www.nice.org.uk/Media/Default/about/what-we-do/research-and-development/Social-Value-Judgements-principles-for-the-development-of-NICE-guidance.pdf.

Niëns, L. (2014). Affordability in Health Care: Operationalizations and Applications in Different Contexts.

Nowotny, H. (2003). Democratising Expertise and Socially Robust Knowledge. *Science and public policy, 30*(3), 151-156.

Oliver, A., Mossialos, E., & Robinson, R. (2004). Health Technology Assessment and Its Influence on Health-Care Priority Setting. *International Journal of Technology Assessment in Health Care, 20*(1), 1-10. doi: 10.1017/S026646230400073X

Patera, N., & Wild, C. (2014). Assessment – Appraisal – Decision. Lbi-Hta Decision Support Document Nr.: 72. Wien: Ludwig Boltzmann Institut für Health Technology Assessment.

Ragin, C. C. (2004). Turning the Tables: How Case-Oriented Research Challenges. *Rethinking social inquiry: Diverse tools, shared standards, 123*.

Rip, A. (1985). Experts in Public Arenas. *Regulating industrial risks: Science, hazards and public protection*, 94-110.

Rip, A. (1986). Controversies as Informal Technology a Ssessment. *Knowledge, 8*(2), 349-371.

Rip, A. (1992). Expert Advice and Pragmatic Rationality. *The Culture and power of knowledge: inquiries into contemporary societies*, 363.

Russell, J. (2017). *The Rationality of Rationing: A Rhetorical Policy Analysis of Deliberations About Resource Allocation in the Nhs.* University of Oxford, UK.

Russell, J., & Greenhalgh, T. (2014). Being 'Rational'and Being 'Human': How National Health Service Rationing Decisions Are Constructed as Rational by Resource Allocation Panels. *Health:, 18*(5), 441-457.

Sayer, A. (2011). *Why Things Matter to People: Social Science, Values and Ethical Life*: Cambridge University Press.

Shah, K. K., Cookson, R., Culyer, A. J., & Littlejohns, P. (2013). Nice's Social Value Judgements About Equity in Health and Health Care. *Health Economics, Policy and Law, 8*(2), 145-165.

Singer, P. A., Martin, D., Giacomini Bhatia, M. V., & Purdy, L. (2000). Priority Setting for New Technologies in Medicine: Qualitative Case Study. *Bmj, 321*(7272), 1316-1318. doi: 10.1136/bmj.321.7272.1316

Stolk, E., Brouwer, W., & Busschbach, J. (2002). Rationalising Rationing: Economic and Other Considerations in the Debate About Funding of Viagra. *Health Policy, 59*(1), 53-63.

Stolk, E., & Poley, M. (2005). Criteria for Determining a Basic Health Services Package. *The European Journal of Health Economics, 6*(1), 2-7.

Wallenburg, I., Weggelaar, A. M., & Bal, R. (2019). Walking the Tightrope: How Rebels "Do" Quality of Care in Healthcare Organizations. *Journal of Health Organization and Management*.

Walley, T. (2007). Health Technology Assessment in England: Assessment and Appraisal. *Medical Journal of Australia, 187*(5), 283.

Zorginstituut Nederland. (2013). Pakketbeheer in De Praktijk Deel 3 (pp. 1-106). Diemen.

Zorginstituut Nederland. (2016). Reglement Adviescommissie Pakket. In Z. Nederland (Ed.). https://www.zorginstituutnederland.nl/over-ons/publicaties/besluit/2016/12/05/reglement-adviescommissie-pakket-zorginstituut-nederland (accessed 11 February 2020).

Zorginstituut Nederland. (2017). Pakketadvies in De Praktijk: Wikken En Wegen Voor Een Rechtvaardig Pakket. Diemen.

S | Summary
Samenvatting

SUMMARY

When exactly is a health care technology *necessary* and how do we determine this on a collective level? Just how far should the collective solidarity stretch? These questions are answered every time a health care coverage decision is made. The Dutch National Health Care Institute (in Dutch: *Zorginstituut Nederland*, in this text: the Institute) is an arm's length body tasked with advising the Minister of Health on these decisions, which specify the contents of the basic benefits basket. The basic benefits basket contains all health care technologies covered by the collective health insurance Dutch citizens are obliged to take out. The Institute uses four specific formalised criteria to formulate their health care coverage advice, namely effectiveness, cost-effectiveness, feasibility, and necessity. This final criterion has been defined or operationalised in different ways over the years since the influential Dunning Committee report on Choices in Health Care first suggested it, nearly 30 years ago. It has, however, also long been considered "non-uniform" and "problematic" by scholars. This dissertation provides a fresh operationalisation of the necessity criterion by answering the question: how is the necessity criterion used in practice?

In this dissertation, I argue that the necessity criterion has, in the past, been operationalised too narrowly, primarily following a well-established tendency at the Institute (and further afield) to guarantee the rationality of decisions through strong adherence to well-explicated decision criteria. This tendency has obscured what this notion of necessity is there to achieve: that an advised decision is not made without explicit reference to the societal context. This reference to the societal context, then, happens in Dutch coverage decision-making practice by using necessity *argumentations*, which are the form the necessity criterion takes in practice. These argumentations, I show, are far broader than the formalised tenets of the necessity criterion as visible in the most recent instalments of relevant policy documentation.

The **Introduction** gives an overview of three cornerstones of this study: first, the most important policy reports, starting with the Dunning Committee report; second, the primarily inductive Health Services Research (HSR) literatures on coverage decision making; and third, the relevant Science and Technology Studies (STS) literatures on expert advice. The policy reports that have dealt with the necessity criterion highlight two things: first, the necessity criterion has been operationalised in many ways, but over the years it seems to have moved from a broad definition to a highly specific one, with elements of severity of illness and reference to the context, whether societal or institutional in terms of insurance, relatively permanent. Second, the location of using the necessity criterion has almost always been in deliberation, and specifically, in the appraisal phase (which is generally defined as a setting for contextualisation or societal weighing and as following the assessment phase, where the scientific input regarding, e.g., effectiveness and cost-effectiveness is established).

The section on HSR zooms in on how health care coverage decisions are made in practice, generally focusing on the deliberative setting. It describes how the field first exhibited an appreciation of making these decisions *not* according to general rules or criteria, as they were seen as too rigid.

Instead, decision makers' personal preferences were described as vital for good decisions. This has evolved to showing expertise in such decisions as doing both-and; the application of criteria, which often comprised using scientific knowledge, and following procedures is combined with local or contextual knowledge, displaying human-ness. Combining these elements well is described as displaying pragmatic rationality, whereby the expertise needed to do so guarantees the quality of the decision.

The STS literatures referenced do not focus on the deliberations only but showcase the role such deliberations and decisions may play in a wider political decision-making process. Here, also, pragmatic rationalism is referenced, but in this case, it is defined as being able to produce a robust outcome in certain contexts. Robustness is described in terms of the ability to withstand pressure 'out there', but also as deriving from substantive input wider and more varied than technical expertise. This focus on what happens to a decision afterwards highlights the importance of a good decision justification or rationale.

Chapter 1 starts with providing an overview of twenty necessity argumentation types derived from scholarly literatures through a realist review. Such a review aims to provide insight into what works, for whom, and in what context. We use it to identify these twenty argumentation types (which include but are not limited to, the morbidity and need as experienced by the patient, availability of alternative treatments, the financial cost per individual patient, and compassion with vulnerable groups such as children or small numbers of patients generally) and examine in what context and by whom they are used to argue in favour of, or against, coverage of the health care technology in question. We conclude that the context wherein a necessity argumentation is used affects its use and outcome in several ways. The use of necessity argumentations thus depends on the outcome of the decision (different argumentation types are used to argue in favour of or against coverage); the person who is arguing in favour of or against coverage (patients, authors, members of the public, and the media use a wider variety of argumentation types than policy makers and insurance companies, and when a member of the public or a patient sets the decision on the agenda the outcome is more likely to be positive); and sometimes on the country where it is used (some countries appear to have very specific argumentation type preferences).

In **Chapter 2**, my co-authors and I report on a cross-country comparison on the use of contextual factors (i.e., necessity argumentations) between Belgium, England, Germany, and the Netherlands. We follow Asdal and Moser in defining contextual factors as situation-specific and compare the countries' use of contextual factors generally based on interviews, and the documentation reporting the justification/rationale of four decisions taken in all four countries. Most interestingly, all four countries do use contextual factors thus operationalised, and the relevant factors are established in deliberation, *around the table*. The Netherlands employs a relatively wide variety of factors and has a high level of formalisation of the necessity criterion compared to Belgium, England, and Germany. Notably, similar contextual factor patterns led to similar decisions across the countries, suggesting patterning of argumentations as an interesting avenue for further research. A significant number of decisions lacked a public justification, raising questions on legitimacy.

Chapter 3 homes in on what happens around the table in the Netherlands, specifically the appraisal phase of four distinct health care coverage decisions, namely eculizumab, front teeth replacement therapy, maternity care, and paracetamol-vitamin D tablets. The appraisal committee (in Dutch: *Adviescommissie Pakket*, *ACP*) is tasked with the 'societal weighing' of these decisions and regarded as a primary place for establishing, or 'constructing', necessity. Through analysis of observations and audio-recordings of appraisal meetings, enriched by interviews with committee members and Institute employees, we show how necessity is constructed in the four phases of appraisal. These four phases are first, the contributions of the Institute employee; second, the contributions of patients and/or their representatives; third, deliberations of the committee; and fourth, taking the final decision. These phases highlight four ways of constructing necessity: 1) allowing explicit criteria to steer the process; 2) allowing patient (representative) contributions to challenge the process; 3) bringing in new argumentations and weaving them together; 4) formulating recommendations. These correspond to four elements of societal weighing expertise, of which the latter two are most distinct, as they show how argumentations from society are actively used and how the decision outcome is embedded in society by the committee.

Chapter 4 derives a three-step model for making robust coverage decisions from Science and Technology Studies literatures on controversy and re-examines some of the data from the previous chapter to illustrate this model. The model conceptualises (advised) decisions as networks as elements and comprises three steps: 1) identifying elements, which include findings, arguments, and values; 2) designing networks of elements, which entails articulating links, broadening the scope of networks, and sometimes black-boxing links; and 3) testing these networks and choosing one. This yields a clear additional rationale for engaging patients and members of the public (as they might contribute different and different types of elements, positively affecting decision network strength) as well as an appreciation of the tacit expertise brought to bear by the committee.

The **Discussion** showcases the contribution of this study by first, summarising how the necessity criterion is used in practice, thus answering the research question, and second, by specifying how societal weighing is achieved through contextualisation practices and delineating how this contributes to HSR and STS literatures.

The necessity criterion takes the form of argumentations in favour of, or against, coverage of the health care technology under consideration, and their perceived validity, and hence their selection differs per case. Many countries use necessity argumentations in coverage decisions, but they generally have formalised these argumentations to differing degrees, with the Netherlands serving as a country with a relatively high level of formalisation. Necessity argumentations are generally established in patterns in a deliberative setting. The Dutch appraisal thus serves as an interesting site for further study, and this is where the societal weighing is achieved.

Societal weighing entails first, 'bringing the outside in', and second, 'placing the decision back outside'. The outside is actively brought in through the presence of patients, their representatives, and/or other stakeholders, who are here for this decision only but steer and guide the deliberations in vital ways and through committee members referencing outside sources such as newspapers

and adding these considerations to decisions. The decision is placed back outside through the way the committee carefully brings lots of argumentations together with producing a robust, that is, able to withstand pressure in outside settings, decision in mind, and through the addition of recommendations relevant for specific outside settings. This adds to HSR and STS literatures as it showcases what makes societal weighing distinct, namely the contextualisation practices, through which the outside, that is, the relevant contexts, affects the processes, contents, and outcomes of the deliberations. This results in different 'types' of societal weighing that may happen around the table, which is not problematic but in fact guarantees the quality of the decision. The deliberations, decisions, and justifications or rationales are of high quality as they display this specific set of contextualisation practices: bringing the outside in and placing the decision back outside in a careful, appropriate, and well-legitimated manner, sensitive to the specific outside contexts that are considered relevant.

I use the metaphor of *weaving necessity* to evoke the art of the decision-making process, wherein the many, case-specific, and patterned necessity argumentations are brought together to help delineate the extent of our collective solidarity appropriately and robustly.

SAMENVATTING

Wanneer is een zorgtechnologie precies noodzakelijk en hoe bepalen we dit op collectief niveau? Hoe ver moet de collectieve solidariteit precies gaan? Deze vragen worden beantwoord op het moment dat er een beslissing wordt genomen over de vergoeding van gezondheidszorg. Zorginstituut Nederland (in deze tekst: het Instituut) is een overheidsinstantie belast met het adviseren van de Minister van Volksgezondheid, Welzijn en Sport over deze beslissingen, waarmee de inhoud van het basispakket van de zorgverzekering wordt gespecificeerd. Het basispakket bevat alle zorgtechnologieën die worden gedekt door de collectieve basisverzekering, die alle Nederlandse burgers verplicht zijn af te sluiten. Het Instituut gebruikt vier specifieke, geformaliseerde criteria om zulk vergoedingsadvies te formuleren, namelijk effectiviteit, kosteneffectiviteit, uitvoerbaarheid en noodzakelijkheid. Sinds het bijna 30 jaar geleden voor het eerst werd voorgesteld in het invloedrijke rapport van de commissie-Dunning, '*Kiezen En Delen, Rapport van de Commissie Keuzen in de Zorg*', is het laatste criterium, noodzakelijkheid, op verschillende manieren gedefinieerd en geoperationaliseerd. Het noodzakelijkheidscriterium wordt echter ook al lang door wetenschappers als "niet-uniform" en "problematisch" beschouwd. In dit proefschrift bied ik een nieuwe operationalisering van dit criterium door de volgende vraag te beantwoorden: hoe wordt het noodzakelijkheidscriterium in de praktijk gebruikt?

In dit proefschrift beargumenteer ik dat het noodzakelijkheidscriterium in het verleden te krap is geoperationaliseerd. De voornaamste reden hiervoor is de sterke neiging bij het Instituut (en daarbuiten) om de rationaliteit van beslissingen te garanderen door middel van rechtlijnig gebruik van duidelijk omschreven beslissingscriteria. Deze neiging verbloemde wat deze noodzakelijkheidsnotie in staat is te bewerkstelligen, namelijk dat een geadviseerde beslissing niet wordt genomen zonder te verwijzen naar de maatschappelijke context. Het noodzakelijkheidscriterium neemt in de praktijk de vorm aan van argumentenen deze argumenten worden gebruikt in de Nederlandse beslissingspraktijk voor deze verwijzing naar de maatschappelijke context. Ik laat zien dat deze argumenten veel breder zijn dan de geformaliseerde principes van het noodzakelijkheidscriterium zoals beschreven in de meest recente relevante beleidsdocumentatie.

De **Inleiding** geeft een overzicht van de drie pijlers van deze studie: ten eerste, de belangrijkste beleidsdocumentatie sinds het rapport van de commissie-Dunning; ten tweede, elevante literatuur over Gezondheidswetenschappen (*Health Services Research*, HSR) over vergoedingsbesluitvorming; en ten derde de relevante Wetenschaps- en Technologiestudies (*Science and Technology Studies*, STS) over advies van deskundigen. De beleidsdocumentatie waarin het noodzakelijkheidscriterium is behandeld benadrukt twee dingen: ten eerste is het noodzakelijkheidscriterium op vele manieren geoperationaliseerd, maar door de jaren heen lijkt de overgang te zijn gemaakt van een relatief brede definitie naar een specifiekere Hierin lijken elementen van ziektelast en verwijzing naar de context, de maatschappelijke context dan wel de geïnstitutionaliseerde context in termen van verzekering, relatief permanent. Ten tweede is het gebruik van het noodzakelijkheidscriterium bijna altijd gelokaliseerd in deliberatie. in de *appraisal* fase om precies te zijn, die over het algemeen

wordt gedefinieerd als de setting voor contextualisering of maatschappelijke weging en als volgend op de *assessment* fase, waar de wetenschappelijke input, onder andere de effectiviteit en kosteneffectiviteit, wordt vastgesteld.

Het gedeelte over HSR zoomt in op hoe beslissingen over de vergoeding van de gezondheidszorg in de praktijk worden genomen, waarbij de nadruk meestal ligt op de setting waarin de deliberatie plaatsvindt. Ik beschrijf daarin hoe het veld startte vanuit een waardering voor het nemen van zulke beslissingen zonder algemene regels of criteria te volgen, omdat deze als te rigide werden beschouwd. In plaats daarvan werden de persoonlijke voorkeuren van besluitvormers beschreven als essentieel voor goede beslissingen. Dit is geëvolueerd naar het beschrijven van expertise in beslissingen die beide dingen doen, namelijk zowel het toepassen van criteria, wat vaak het gebruik van wetenschappelijke kennis en het volgen van procedures omvat, als het in acht nemen van lokale of contextuele kennis, waarin de menselijkheid van de commissie wordt getoond. Het goed combineren van deze elementen wordt omschreven als het demonstreren van pragmatische rationaliteit, waarbij de benodigde expertise de kwaliteit van de beslissing garandeert.

De STS-literatuur die ik behandel richt zich niet alleen op de deliberatie, maar onderstreept vooral de rol die dergelijke beraadslagingen en beslissingen kunnen spelen in een breder politiek besluitvormingsproces. Ook hier wordt de term pragmatisch rationalisme gebruikt, maar in dit geval wordt het gedefinieerd als het in staat zijn om een robuust resultaat te produceren in bepaalde contexten. Robuustheid wordt beschreven in termen van het vermogen om druk van 'buiten' te weerstaan, maar wordt ook afgeleid uit inhoudelijke input, die breder en gevarieerder is dan alleen technische expertise. Deze focus op wat er met een beslissing gebeurt als deze eenmaal is genomen onderstreept het belang van een goede motivering van een beslissing.

Hoofdstuk 1 begint met een overzicht van twintig noodzakelijkheidsargumentatietypen afgeleid van wetenschappelijke literatuur door middel van een *realist review*. Dit type review is bedoeld om inzicht te geven in wat werkt, voor wie en in welke context. We hebben deze methode gebruikt om deze twintig argumentatietypes (waaronder de morbiditeit en behoeftes van de patiënt, de beschikbaarheid van alternatieve behandelingen, de financiële kosten per individuele patiënt en compassie met kwetsbare groepen zoals kinderen of kleine aantallen patiënten) te identificeren. We onderzoeken vervolgens in welke context en door wie ze worden gebruikt om te pleiten voor, of tegen, vergoeding van de technologie in kwestie. We concluderen dat de context waarin een noodzakelijkheidargumentatie wordt gebruikt op verschillende manieren invloed heeft op het gebruik en de uitkomst van de beslissing. Het gebruik van noodzakelijkheidsargumenten hangt dus af van de uitkomst van de beslissing (verschillende argumentatietypes worden gebruikt om voor of tegen vergoeding te pleiten); van de persoon (patiënten, auteurs, burgers en de media gebruiken een breder scala aan argumentatietypes dan beleidsmakers en verzekeringsmaatschappijen, en wanneer een burger of patiënt het besluit op de agenda plaatst, is het resultaat eerder positief); en soms van het land waar het wordt gebruikt (sommige landen lijken zeer specifieke voorkeuren voor argumentatietypes te hebben).

In **Hoofdstuk 2** rapporteren mijn coauteurs en ik over het gebruik van contextuele factoren (d.w.z. noodzakelijkheidargumenten) in een vergelijking tussen België, Duitsland, Engeland en Nederland. We volgen Asdal en Moser in het definiëren van contextuele factoren als situatiespecifiek. We vergelijken het gebruik van contextuele factoren in het algemeen in deze landen op basis van interviews, en specifiek op basis van documentatie die de motivering van vier specifieke beslissingen in alle vier de landen rapporteert. Het meest interessante is dat alle landen dergelijke contextuele factoren gebruiken, en de relevante factoren worden in overleg, rond de tafel, vastgesteld. Nederland hanteert een relatief grote verscheidenheid aan factoren en kent een hoge mate van formalisering van het noodzakelijkheidscriterium in vergelijking met België, Duitsland, en Engeland. Vergelijkbare contextuele factorpatronen leiden tot vergelijkbare beslissingen in de verschillende landen, wat suggereert dat argumentatie in patronen wordt gevormd, wat interessant zou kunnen zijn voor verder onderzoek. Een aanzienlijk aantal beslissingen bevatte geen openbare motivering, wat vragen oproept over de legitimiteit van deze beslissingen.

In **Hoofdstuk 3** wordt ingegaan op wat er in Nederland om de tafel gebeurt, met name in de *appraisal* fase, in vier verschillende vergoedingsbesluiten, namelijk eculizumab, fronttandvervanging, kraamzorg en paracetamol-vitamine D-tabletten. De Adviescommissie Pakket (ACP, in deze tekst: *appraisal* commissie) is belast met de 'maatschappelijke weging' van deze beslissingen en wordt beschouwd als de primaire plaats voor het vaststellen of 'construeren' van noodzakelijkheid. Door analyse van observaties en audio-opnames van commissievergaderingen, verrijkt met behulp van interviews met commissieleden en medewerkers van het Instituut, laten we zien hoe noodzakelijkheid wordt geconstrueerd in vier fasen van de *appraisal* vergadering. Deze vier fasen zijn: ten eerste, de bijdragen van de medewerker van het Instituut; ten tweede, de bijdragen van patiënten en/of hun vertegenwoordigers; ten derde, de beraadslagingen van de commissie; en ten vierde, het nemen van de definitieve beslissing. Deze fasen belichten vier manieren om noodzakelijkheid te construeren: 1) toelaten dat expliciete criteria het proces sturen; 2) toelaten dat bijdragen van patiënten of hun vertegenwoordigers het proces uitdagen; 3) nieuwe argumentaties inbrengen en deze combineren; en 4) formuleren van aanbevelingen. Deze vier manieren komen overeen met vier elementen van deskundigheid op het gebied van maatschappelijke weging, waarvan de laatste twee het meest specifiek zijn omdat ze laten zien hoe argumentaties uit de samenleving actief worden gebruikt en laten zien hoe het besluit door de commissie in de samenleving wordt ingebed.

In **Hoofdstuk 4** beschrijven we een driestappenmodel voor het nemen van robuuste vergoedingsbesluiten op basis van STS-literatuur over controverse, geïllustreerd met behulp van data uit het vorige hoofdstuk. Wij conceptualiseren (geadviseerde) beslissingen als netwerken van elementen en het model om deze te bereiken bestaat uit de volgende drie stappen: 1) het identificeren van elementen, waaronder bevindingen, argumenten en waarden; 2) het ontwerpen van netwerken van elementen, waaronder verbindingen tussen elementen leggen, het verbreden van de reikwijdte van netwerken, en soms het 'black-boxen' van verbindingen; en 3) deze netwerken testen en er een kiezen. Dit levert een duidelijke (aanvullende) reden op om patiënten en burgers te betrekken (omdat ze verschillende en verschillende soorten elementen kunnen bijdragen, die de sterkte van

het beslissingsnetwerk positief beïnvloeden), alsook een waardering van de expertise van de commissie.

In de **Discussie** expliciteer ik de bijdrage van deze studie door eerst samen te vatten hoe het noodzakelijkheidscriterium in de praktijk wordt gebruikt en daarmee de onderzoeksvraag te beantwoorden, en ten tweede door te specificeren hoe een maatschappelijke weging wordt bereikt door middel van 'contextualiseringspraktijken', en duidelijk te maken hoe dit bijdraagt aan HSR- en STS-literatuur.

Het noodzakelijkheidscriterium neemt de vorm aan van argumenten voor of tegen vergoeding van een bepaalde medische technologie. De geldigheid van deze argumenten, en dus hun selectie, verschilt per geval. Veel landen gebruiken noodzakelijkheidsargumentaties in vergoedingsbesluiten, maar ze hebben deze argumentaties in verschillende mate geformaliseerd. Nederland is duidelijk een land met een relatief hoog formaliseringsniveau. In de setting van de deliberatie worden noodzakelijkheidargumenten meestal in patronen gebruikt. De Nederlandse *appraisal* dient dus als een interessante plek voor verder onderzoek omdat juist hier de maatschappelijke weging plaatsvindt.

Maatschappelijk wegen omvat ten eerste, 'het buiten binnen brengen', en ten tweede 'de beslissing terug naar buiten brengen'. Het buiten wordt actief binnengebracht door de aanwezigheid van patiënten, hun vertegenwoordigers en/of andere stakeholders, die alleen voor deze beslissing aanwezig zijn en de beraadslagingen op cruciale manieren sturen en begeleiden, en door de commissieleden die gebruik maken van externe bronnen zoals kranten en deze overwegingen aan beslissingen toevoegen. De beslissing wordt terug naar buiten gebracht door de manier waarop de commissie zorgvuldig een veelheid aan argumentaties combineert met als doel het produceren van een robuuste beslissing, dat wil zeggen, een beslissing die bestand tegen druk van situaties buiten. Hetzelfde gebeurt door de toevoeging van aanbevelingen die relevant zijn voor specifieke externe instellingen. Dit draagt bij aan HSR- en STS-literatuur omdat het laat zien wat maatschappelijk wegen onderscheidt, namelijk de contextualiseringspraktijken, waardoor het buiten, dat wil zeggen de relevante contexten, de processen, inhoud en resultaten van de beraadslagingen beïnvloedt. Dit resulteert in verschillende 'soorten' maatschappelijk wegen die rond de tafel kunnen plaatsvinden. Dit is echter niet problematisch, maar garandeert in feite de kwaliteit van de beslissing. De beraadslagingen, beslissingen en motivaties zijn van hoge kwaliteit omdat ze deze specifieke set van contextualiseringspraktijken laten zien: het buiten binnenbrengen en de beslissing terug naar buiten brengen op een zorgvuldige, passende en goed onderbouwde manier die gevoelig is voor de specifieke contexten die als relevant worden beschouwd.

Ik gebruik de metafoor van het weven van noodzakelijkheid om de kunst van het besluitvormingsproces op te roepen, waarbij de vele, casus-specifieke noodzakelijkheidargumenten in een patroon worden samengebracht om onze collectieve solidariteit op een passende en robuuste manier te helpen afbakenen.

Portfolio
Bio

PORTFOLIO

Peer-reviewed publications

Kleinhout-Vliek, T., de Bont, A., & Boer, B. (2017). **The bare necessities? A realist review of necessity argumentations used in health care coverage decisions**. Health Policy, 121(7), 731-744.

Signatory to: Newton, P. N., & Bond, K. C. (2019). **Global access to quality-assured medical products: the Oxford Statement and call to action.** The Lancet Global Health, 7(12), e1609-e1611.

Kleinhout-Vliek, T., de Bont, A., Boysen, M., Perleth, M., van der Veen, R., Zwaap, J., Boer, B. (2020). **Around the Tables – Contextual Factors in Healthcare Coverage Decisions across Western Europe.** International Journal of Health Policy and Management, 9(9), 390-402.

Kleinhout-Vliek, T., de Bont, A., & Boer, B. (2020). **Necessity under Construction – societal weighing rationality in the appraisal of health care technologies –** *accepted to Journal of Health Economics, Policy, and Law*

Kleinhout-Vliek, T., de Bont, A., & Boer, B. **Arguments in networks – a conceptual model for robust health care coverage decisions –** *under review at Sociology of Health & Illness*

Kootstra, J. & Kleinhout-Vliek, T. **Implementing pharmaceutical track-and-trace systems – a realist review** – *submitted to BMJ Global Health*

M. Felder, T. Kleinhout-Vliek*, M. Stevens*, & A. de Bont. **Design thinking experiments: an ethnographic case study –** *submitted*

T.H. Kleinhout-Vliek*, M.B. de Graaff*, H.M. van de Bovenkamp & R.A. Bal. **Patient and Public Involvement and Engagement in health care decision making: a research agenda –** *to be submitted*

** equal contributors*

Other publications

Kleinhout-Vliek, T. (2017). **You can't say fairer than that!** European Association of Centres of Medical Ethics Newsletter December 2017 – *invited contribution*

M.B. de Graaff, H.M. van de Bovenkamp, M.G.E. Crombach, T.H. Vliek & R.A. Bal (2018). **Burgers en patiënten betrekken bij beslissingen: voor altijd in de kinderschoenen?** Rotterdam: ESHPM

Presentations

Wait – is this really necessary care? Using necessity as a criterion for priority setting decisions
International Society on Priorities in Health 11th biennial meeting
University of Birmingham
7 September 2016

Bittere noodzaak of overbodige luxe? En hoe bepalen we dat? *invited talk*
Zorginstituut Nederland
16 May 2017

Societal weighing in deliberative health care coverage decisions
European Association of Centres of Medical Ethics
'Justice in Health Care: Values in Conflict'
Institut Borja de Bioètica-URL, Barcelona
8 September 2017

Noodzakelijkheid, een lastig criterium *movie*
1e Wetenschapsdag Zorginstituut Nederland
5 October 2017

Noodzakelijkheid bepaal je om tafel *invited talk*
Symposium 'Pakketbeheer: kennis over keuzes'
Erasmus Universiteit Rotterdam
23 March 2018

Noodzakelijkheidscriterium in pakketbeslissingen *invited talk*
Zorginstituutforum, Zorginstituut Nederland
20 June 2019

Around the tables – reflections on doing 'bottom-up' research
Research colloquium Critical and Interpretative Public Administration
Radboud Universiteit Nijmegen
21 June 2019

Attended summer schools, workshops, conferences

Summer School 'Health Law and Ethics'
The Erasmus Observatory on Health Law, Rotterdam
1-10 July 2015

Summer School 'Politics of Science, Technology, and STS'
Netherlands Graduate Research School of Science, Technology and Modern Culture, Nijmegen
24-28 August 2015

Conference 'Individuals, Public Interests and Public Goods: What is the Contribution of Bioethics?'
World Congress of the International Association of Bioethics, Edinburgh
15-17 June 2016

Workshop 'Prototyping Intervention' *selected participant*
Technical University of Munich
24-25 October 2017

Conference 'Medicine Quality and Public Health'
University of Oxford
25-28 September 2018

Summer School 'Pharmaceutical Policy Analysis'
Utrecht University
15-19 July 2019

Courses

A successful doctoral track
September 2015

Group dynamics
September 2015

Atlas.ti
September 2015

Scientific integrity day
October 2015

Doing the literature review
Winter 2015-2016

Academic writing in English
Spring 2016

How to get your article published
Spring 2017

Shut up and write
November 2017

Basic didactics
Winter 2017-2018

Analytic storytelling
October 2019

BIO

Trijntje Hendrika (Tineke) Kleinhout-Vliek (1986) is a qualitative researcher interested in the governance of medicine access. She was originally trained in biomedical research (BSc and MSc Molecular Life Sciences, Wageningen University & Research, internship at the University of Edinburgh). Tineke found her groove during three months' fieldwork in India for her second MSc (Development & Rural Innovation, Wageningen University & Research, *cum laude*), investigating how India managed to get rid of the patent on an expensive leukaemia medicine despite international patenting legislation. This dissertation continues her interest in the governance of medicine access by examining the appraisal phase of Dutch health care coverage decisions, outlining how a societal weighing is achieved therein through contextualisation practices, utilising insights from Science and Technology Studies literatures.

Tineke is currently working at the Copernicus Institute of Sustainable Development, Utrecht University, as part of an international research project on social pharmaceutical innovation to benefit unmet medical needs. She is fascinated by novel ideas and enjoys discussing them, whether over a cup of coffee or in front of a lecture hall. Tineke lives in The Hague with her husband Tom and their two wee girls.

All health care technologies covered through the basic benefits basket of the Dutch collective health insurance should be effective, provide value for money, and be necessary. This last criterion determines the extent of the collective solidarity. But how to establish whether a health care technology is necessary?

Through examining qualitative cases from the Netherlands and further afield, Tineke Kleinhout-Vliek describes how necessity is established in health care coverage decision-making practice. Manifold argumentations pertaining to necessity exist and are used by the Dutch National Health Care Institute and its appraisal committee (and their equivalents in several Western European countries) when deliberating on and formulating advised health care coverage decisions. Kleinhout-Vliek details the varied but patterned use of these necessity argumentations in cases ranging from highly expensive medicines to paracetamol tablets and maternity care.

Such varied use of necessity argumentations may seem problematic, as variation would indicate inconsistency and potential inequity between decisions. Instead, Kleinhout-Vliek shows that these variable necessity argumentations are used for setting health care coverage decisions in context, that is, for contextualising them. As the relevant contexts differ per health care technology appraised, the relevant necessity argumentations vary likewise. Well-contextualised health care coverage decisions contain necessity argumentations that are carefully woven together. Weaving necessity makes these decisions robust: able to withstand pressure outside the decision-making setting.