Responsible Prediction under Critical Uncertainty

AN EPISTEMIC ANALYSIS OF NEUROPROGNOSTIC
INNOVATION PRACTICES AFTER CARDIAC ARREST

Mayli Mertens

Responsible Prediction under Critical Uncertainty

AN EPISTEMIC ANALYSIS OF NEUROPROGNOSTIC
INNOVATION PRACTICES AFTER CARDIAC ARREST

Mayli Mertens

RESPONSIBLE PREDICTION UNDER CRITICAL UNCERTAINTY

AN EPISTEMIC ANALYSIS OF NEUROPROGNOSTIC
INNOVATION PRACTICES AFTER CARDIAC ARREST

DISSERTATION

to obtain

the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof.dr.ir. A. Veldkamp,

on account of the decision of the Doctorate Board,
to be publicly defended

on Wednesday the 29th of September 2021 at 12.45 hours

by

Mayli Mertens

born on the 9th of August 1982
in Edegem, Belgium

This thesis has been approved by

The promotor:
prof. dr. M. Boenink

The co-promotor:
prof. dr. P.A.E. Brey

Cover and Illustrations: Anya Lauri Illustrations
Printed by: Ipskamp Printing
Lay-out: Caroline Axelblom
ISBN: 978-90-365-5234-9
DOI: 10.3990/1.9789036552349

© 2021 by Mayli Mertens. Enschede, The Netherlands. All rights reserved.
No parts of this thesis may be reproduced, stored in a retrieval system
or transmitted in any form or by any means without permission of the
author. Alle rechten voorbehouden. Niets uit deze uitgave mag worden
vermenigvuldigd, in enige vorm of op enige wijze, zonder voorafgaande
schriftelijke toestemming van de auteur.

EXAMINATION COMMITTEE:

Chairman/secretary prof.dr. Th.A.J. Toonen

Supervisors prof. dr. M. Boenink (University of Twente)
 prof. dr. P.A.E. Brey (University of Twente)

Members prof. dr. M.E.H. Ong (Duke-NUS)
 dr. J. Mesman (Maastricht University)
 dr. E.O.J. Kompanje (Erasmus MC)
 prof. dr. ir. M. Boon (University of Twente)
 prof. dr. C. Aydin (University of Twente)

This work was funded by the Dutch Organization for Scientific Research (NWO), under its programme Responsible Innovation (NWO-MVI) [grant number 313 99 309], with co-funding from Twente Medical Systems international (TMSi), Clinical Science Systems (CSS), and the Dutch Brain Foundation (Hersenstichting).

TABLE OF CONTENTS

Preface	xi
● Introduction responsible innovation in neuroprognostication after cardiac arrest.....	1
● CHAPTER 1 Liminal Innovation Practices: questioning three common assumptions in responsible innovation	15
1. Introduction	16
2. Case Study : RI in prognosis of postanoxic coma	16
3. Three assumptions in RI	18
3.1 Background.....	18
3.2 Assumption 1: emerging technologies, radically novel and highly unpredictable.....	19
3.3 Assumption 2: impacting innovation trajectories early on.....	20
3.4 Assumption 3: the need for anticipation.....	20
4. Innovations in clinical settings: questioning the assumptions.....	21
4.1 Questioning assumption 1: technology meets incremental innovation..	21
4.2 Questioning assumption 2: a continuous lemniscate of innovation.....	23
4.3 Questioning assumption 3: the need for anticipation in clinical settings.....	24
5. Betwixt and between: liminal innovation practices in clinical settings.....	26
5.1 Between research and care, between experimentation and implementation.....	27
5.2 Between previous and future innovation.....	28
5.3 Not yet used, yet not 'not used'.....	29
6. Conclusion: Meaning for RI approaches.....	31
7. Discussion: Different contexts require different approaches	33
● CHAPTER 2 Braving Uncertainty After Cardiac Arrest: how two questions without definite answers shape a clinical dilemma.....	37
1. Introduction: Aim and Motivation.....	38
2. Methodology.....	39
3. Uncertainty and the struggle to predict and control.....	41

4. Making decisions in the face of uncertainty: the dilemma.....	43
4.1 Choosing to withdraw treatment.....	44
4.2 Choosing to continue treatment.....	45
4.3 Delay or advance decision-making: extended and reduced windows of opportunity.....	48
5. Braving uncertainty: avoiding the greatest harm.....	50
5.1 The suffering of the patient versus their life lost.....	51
5.2 Emotional and psychological impact for next of kin and caregivers.....	53
5.3 Societal cost and distributive justice.....	57
6. Conclusion.....	61

● CHAPTER 3

Predicting Medical Futility After Cardiac Arrest:

pulling apart physiological and normative uncertainty.....	65
--	----

1. Introduction.....	66
2. Judgments of medical futility.....	67
3. Uncertainty and neuroprognostic tools.....	69
4. Evaluating normative futility.....	71
4.1. Views on quality of life and where to draw the line are diverse.....	72
4.2. Views on quality of life and where to draw the line are volatile.....	76
4.3. Views on quality of life and where to draw the line are context dependent.....	79
5. Pulling apart physiological and normative uncertainty.....	82
5.1. Two distinct uncertainties requiring two distinct predictions.....	83
5.2. Who predicts what?.....	84
6. Conclusion.....	86

● CHAPTER 4

Chasing Certainty After Cardiac Arrest:

towards responsible implementation of cEEG for neuroprognostication.....	89
---	----

1. Introduction.....	90
2. Post-anoxic coma, prognostication, and continuous EEG-monitoring.....	92
3. Theoretical background: technology, its performativity and architecture.....	93
4. Methodology.....	94
5. cEEG in practice: challenges and unintended effects.....	96
5.1. Facilitating interpretation.....	96
5.1.1. Identifying patterns.....	96
5.1.2. Classifying EEG-measurements.....	97
5.1.3. Linking measurements to outcomes.....	99
5.2. Facilitating decision-making.....	100

5.2.1. Responding to 'poor' outcomes.....	101
5.2.2. Responding to 'good' outcome.....	102
5.2.3. Responding to uncertain outcomes.....	103
5.3. Facilitating communication.....	105
5.3.1. Visualization of results.....	106
5.3.2. Timing of disclosure of prognostic outcome.....	110
6. Summing up and looking forward.....	111

● CHAPTER 5

Preferences for prognostication of coma after cardiac arrest: the Dutch societal perspective on physiological and normative futility.....

115

1. Introduction.....	116
2. Materials and Methods.....	117
2.1. Study Design.....	117
2.2. Survey Instrument.....	118
2.3. Study Population.....	120
2.4. Statistical Analysis.....	120
3. Results.....	121
3.1. Background Characteristics.....	121
3.2. Perspectives on quality of life.....	122
3.3. Preference for prognostic information.....	123
3.4. Attitude toward withdrawal of life-sustaining treatment.....	125
3.5. Involvement in decision making.....	126
4. Discussion.....	126
5. Strengths and limitations.....	127
6. Conclusion.....	128

● CHAPTER 6

When hidden mistakes impede learning... taking practical and epistemic responsibility for self-fulfilling prophecies.....

131

1. Introduction.....	132
2. Descriptive Account: what are SFPs and how do they work?.....	133
2.1. Credibility as a prediction.....	133
2.2. Employment of the prediction.....	135
2.3. Employment Sensitivity.....	136
2.4. Realization of the prediction.....	138
2.5. Defining the self-fulfilling prophecy.....	140
2.6. Collective SFPs.....	141
3. Normative Analysis: what is problematic about SFPs?.....	143

3.1. Current concerns regarding unintended outcomes and accountability	143
3.2. Taking practical and epistemic responsibility for self-fulfilling prophecies	145
4. Conclusion.....	149

● CHAPTER 7

Self-Fulfilling Prophecies & Responsible Neuroprognostic Innovation:

recognizing invisible physiological and normative prognostic errors.....	153
1. Background: filling the gap	154
2. Facing physiological uncertainty. What is the problem?	157
2.1. First: remaining physiological uncertainty in neuroprognostication	158
2.2. Second: Probabilistic decisions are never error proof.....	159
2.3. Third: SFPs produce no error signal nor reliable positive feedback.....	160
2.4. Fourth: Error signals and reliable feedback are fundamental to learning.....	163
2.5. Finally: Without correct learning we cannot accurately reduce uncertainty.....	164
3. Learning inhibitors iterate mistakes in predictions of physiological futility.....	165
3.1. Established prognostic tests: including past mistakes in blinded research.....	166
3.2. Liminal innovation practices: including current mistakes	167
3.3. Continuous innovation: including new cases in ongoing research	168
4. Facing normative uncertainty. Are we learning when we can?.....	169
4.1. Normative evaluations of outcomes in neuroprognostication.....	170
4.2. Undetected false positives and negatives due to categorisation issues	171
4.3. Self-fulfilling predictions of normative futility and utility.....	173
5. Guidelines: Can we learn from hidden mistakes in the face of physiological and normative uncertainty?	177
6. Conclusion.....	178

● CONCLUSION

Pitfalls and Opportunities for Responsible Neuroprognostic Innovation	181
--	------------

References	192
Acknowledgments.....	211
Dutch Summary.....	221

Preface

When writing my introduction, many advised me to keep it dry, void of narrative or personal anecdotes; to stick to the research questions, what motivated them, and the context that was required for the reader—specifically the committee—to understand them. Those of you who know me, or the woman that raised me, know that I don't take directions well unless I've been given convincing arguments. Context is essential and, often, context is personal.

When I was 6 years old, my younger brother suffocated in his bed. After several resuscitation attempts, he vomited. This response gave my mother hope that he would regain consciousness. He didn't. He died from brain damage due to oxygen deprivation in what I now know is called postanoxic coma—the subject of this research.

In the first year of my doctorate, after spending 7 months doing empirical observations in intensive care, my dad had a brain hemorrhage. While he was in a coma, my older brother, my two sisters, and I had to decide—based on what we knew to be our father's preferences—whether or not to withdraw life-sustaining treatment. In the end, due to a bleed in his brain stem, three of us had to convince my little sister who had just given birth to her second child and had recently also lost her mom (we have a complicated family) to let our father go. The final challenge was that my dad was hospitalized in Thailand and withdrawal of life-sustaining treatment is illegal there.

For those who believe that these kinds of personal background stories and an evocative writing style may fit in a novel but have no place in scientific work, let me share with you the wise advice of multidisciplinary artist, Tim Minchin.

Please don't make the mistake of thinking the arts and sciences are at odds with one another. That is a recent, stupid, and damaging idea. You don't have to be unscientific to make beautiful art, to write beautiful things. If you need proof —Twain, Douglas Adams, Vonnegut, McEwan, Sagan and Shakespeare, Dickens for a start. You don't need to be superstitious to be a poet. You don't need to hate GM technology to care about the beauty of the planet. You don't have to claim a soul to promote compassion. Science is not a body of knowledge nor a belief system. It's just a term which describes human kinds' incremental acquisition of understanding through observation. Science is awesome! The arts and sciences need to work together to improve how knowledge is communicated.

Using the arts to help communicate scientific knowledge does not mean confusing anecdotal and personal experience for evidence. I was so worried that my personal trajectory would incorrectly influence my work that I requested psychological support to make sure I wouldn't get confused or biased. As it turned out, the opposite was true.

First, seeing how the practices in Thailand affected patients, their families, as well as the medical practitioners further confirmed the importance of the multi-contextual, comparative fieldwork I had planned. Second, during the first half year of my empirical work, I observed intensive care practices but I did so mostly through the eyes of the clinicians I followed. They discussed with me their perspective, and their perspectives influenced my interpretation of what I saw. By the time I approached my dad's hospital bed, my siblings reproached me for being unemotional, distant, and—dare I say it—clinical. My work had already given me a professional malformation and seeing the people I love suffer from what had become yet another case study to me, reminded me of what loss really meant. Remembering my own loss from long ago helped me balance out the perspective I had adopted and, when I returned to the intensive care unit, I was able to approach families in the midst of their grief and nevertheless get their consent to participate in a qualitative study. It was thanks to my siblings that my respondents opened up to me in ways they could not to any of my colleagues. It was thanks to the understanding I gained through personal experience that my respondents invited me to funerals and memorial services, years after I interviewed them. While I hope my thesis does justice to their testimonies, it remains first and foremost scientific work.

The goal of this work is to understand whether, and if so, how, a new technological tool should be introduced in neuroprognostic practice. The stories I just told should remind you of the significance that both the questions and the answers will have. While this is a philosophical endeavor, it focuses on practical reality. Whatever the choices we make, whether in national legislation, clinical guidelines, daily hospital practice, or personal judgment, they will impact people's lives... and their demise.

"Deciding whether and when to use a new tool can be one of the most difficult challenges in clinical practice."

Kenneth W. Goodman – Ethics, Medicine, and Information Technology

"The only certainty, whether spoken or not, is that the doctors, nurses, and technicians are fighting not only death, but their own uncertainties as well. In most resuscitations, those uncertainties can be narrowed down to two main questions: Are we doing the right things? and, Should we be doing anything at all?"

Sherwin B. Nuland – How We Die

... to my mother and the one we lost

Introduction:

responsible innovation in neuroprognostication after cardiac arrest

PROGNOSTICATION AFTER CARDIAC ARREST

Out of hospital cardiac arrest (OHCA) is fatal, nine times out of ten (Hassager et al. 2018). Most people die within minutes after their heart stops but even when a reanimation attempt is successful, restarting the heart cannot undo any brain damage caused by oxygen deprivation during the arrest. Therefore, even after return of spontaneous circulation (ROSC), patients commonly remain unresponsive and are generally treated in intensive care for what is called *postanoxic coma*, meaning coma after an oxygen deficit. This coma can last from several days to several weeks and about half of the patients in this state will die from neurological damage (Bro-Jeppesen et al. 2009, Cronberg et al. 2015). Thus, only about half of patients in postanoxic coma survive to hospital discharge with functional neurological health (Hassager et al. 2018) although cognitive impairment in survivors is common (Moulaert 2014).

Within the first day after successful reanimation, the mortality for patients in postanoxic coma is generally due to circulatory failure. However, the main cause of death after the first 24 hours is withdrawal of life sustaining treatment (WLST) following a prognosis of severe neurological damage (Hassager et al. 2018). Reliable neuroprognostication of patients in postanoxic coma is therefore crucial. Without it, decisions on whether or not to continue treatment pose two risks: either withdrawing life-sustaining treatment in patients who may survive with functional recovery or continuing treatment in patients who survive in conditions they would not have consented to (e.g. prolonged disorders of consciousness like vegetative state or minimally conscious state).

Unfortunately, a degree of uncertainty always remains in prognosis (Smith et al. 2013) and neuroprognostication after cardiac arrest is especially precarious. Predicting how much damage the brain suffered is very difficult and the implications for the patient's mental and physical recovery are only really apparent after the patient regains consciousness—if they regain consciousness at all. Better prognostic tools are needed but each tool has its own epistemological and ethical shortcomings. With this research, I sought to identify and analyze these shortcomings so that they can adequately be addressed. This would help ensure responsible innovation of the neuroprognostic practice for patients in coma after cardiac arrest.

Throughout most of this research project, until April 2019, both the Netherlands and the United States (among most other countries) had included only two indicators for poor prognosis in their protocols for neuroprognostication of postanoxic coma (Zandbergen et al. 2011, Taccone et al. 2014). Both indicators rely on testing brainstem reflexes and are still in use today¹. These tests are typically administered seventy-two hours after ROSC or after the patient achieves normothermia, or normal body temperature, in cases where therapeutic hypothermia was applied. Therapeutic hypothermia requires cooling of the body to prevent fever which could otherwise further damage the brain. The existing protocols furthermore recommended gathering data from other prognostic tools and tests². However, these additional tools still needed to be further investigated to establish their predictive value as robust.

The limitation that the existing prognostic protocols posed for predicting outcome after cardiac arrest has been problematic for two reasons. First of all, many of the patients who ultimately have a 'poor outcome' (meaning they pass away or suffer severe neurological damage and disability) still *have* the brainstem reflexes that would otherwise indicate poor outcome³. This means that only a fraction of all patients who have a poor outcome can be identified as such. Therefore, prognostic tools that identify a greater number of poor outcomes, either individually or in aggregate, are needed. Second, the existing tools tell us nothing about good outcomes. All healthcare practitioners can do is actively look for signs that someone will do badly, but there are no real tools available to predict someone will do well. Combined, this leaves about 75% of the patient group with no prognostic information at all.

In order to better predict, during postanoxic coma, what kind of neurological damage patients and their families can expect, other prognostic tools are being investigated. Electroencephalogram monitoring (EEG) has proven to be a promising innovation in that regard (Westhall et al. 2016) with continuous EEG monitoring (cEEG) further improving the prognostic capacity of the tool (Hofmeijer et al. 2015). Even more promising results are obtained when cEEG-based prognosis is combined with machine learning techniques to further explore its predictive potential (Tjepkema-Cloostermans et al. 2017, Nagaraj et al. 2018, Tjepkema-Cloostermans et al. 2019).

¹ Brainstem reflexes can be tested through either clinical examination of pupillary and corneal reflexes or through a somatosensory evoked potentials test (SSEP-test) checking cortical responses.

² Prognostic tools such as electroencephalography (EEG); serum biomarkers like Neuron Specific Enolase (NSE); and neuroimaging such as Computerized Tomography (CT) and Magnetic Resonance Imaging (MRI) including novel sequences such as Diffusion Tensor Imaging (DTI).

³ Meaning that corneal and pupillary reflexes or cortical responses are still present.

These promising advances are not without challenges, however. While improving predictive capacities is expected to better inform decisions on WLST, technological innovations may deepen existing controversies with regard to the quality of life of surviving patients. It may also aggravate tensions between personal values regarding the wishes of the patient and their family and societal concerns such as cost reduction and fair distribution of limited resources. Finally, EEG-based technology enables ultra-early prognostication (Hofmeijer et al. 2015, Ruijter et al. 2019, Tjepkema-Cloostermans et al. 2019). This may exacerbate concerns regarding the accuracy (Geocadin et al. 2012) and timing (Kitzinger & Kitzinger 2013, Gold et al. 2014) of the prognosis and subsequent decision-making. On a methodological level, the development of new neuroprognostic technology requires research within the clinical setting, with actual patients. This impacts both the care of the patients included in the research, as well as the prognosis of future patients, once the new technology is fully implemented.

These multiple challenges are quite complex. Moreover, standards for responsible innovation call for an investigation of potential risks and undesirable consequences of neuroprognostic innovation and an analysis on how best to avoid them. This introduction offers only a provisional problem statement. One important goal of this dissertation is to better identify the problems and articulate them in a more nuanced and detailed way.

The findings and conclusions presented in this dissertation have become ever more relevant for clinical practice, as an increasing number of countries are including cEEG as part of their multimodal prognostic model. In April 2019, the Netherlands included cEEG in their national guidelines for indication of poor outcome, as early as twenty-four and twelve hours⁴ after admission to the ICU (Horn et al. 2019). As of 2021, the new European guidelines also give more weight to the predictive value of EEG in their multimodal prognostication model for patients who remain in coma after 72 hours or more.⁵ (Nolan et al. 2021).

⁴ Following the new Dutch guideline, 'reliable predictors of poor neurological outcome are: iso-electric or suppression-EEG from 12 hours after resuscitation; low voltage EEG from 24 hours; burst-suppression with identical bursts from 24 hours; and GPDs on an iso-electric background pattern from 24 hours' (Horn et al, 2019)

⁵ EEG is now included alongside SSEP and corneal and pupillary reflexes, but also NSE and CT or MRI as indicative of poor outcome, when at least two of the listed test results indicate it, while discordant signs indicating a potentially good outcome should warrant caution.

AIM OF THE STUDY - Research Questions

The purpose of this research is to make sure that neuroprognostic innovation improves prognostic practices for patients in coma after cardiac arrest while diminishing potential downsides to a minimum. Ensuring responsible research, development, and implementation of the innovative technology demands we identify, analyze, and address the potential challenges of neuroprognostic innovation.

The main aim of this research is thus to investigate *how to ensure responsible innovation of prognostication in postanoxic coma*. The novel use of cEEG in prognosis of postanoxic coma serves as my case study as it is the best object of investigation for answering the main research question. Aside from being the most promising prognostic tool for this patient group, it turns out that many of EEGs challenges are challenges that established as well as future prognostic tools face as well.

To fulfill the aim of this study, I first explore *how cEEG-enabled prognosis may affect the current concerns in care for postanoxic patients in coma within the Netherlands*. This requires, a) the identification of current concerns and, b) a way to envision the potential effects of the new technology on those concerns.

Second, based on the findings, I make suggestions on *how to develop the cEEG-based prognostic technology in such a way that it addresses existing concerns and takes into account potentially undesirable effects* so that it can be considered a responsible innovation.

Third, I distil from the case study *what findings can be generalized to neuroprognostic innovation practices for patients in coma and, broader still, to innovation practices in clinical settings*. This is where I focus on the broader ethical and epistemological shortcomings of neuroprognostic innovation for patients in coma.

In the conclusion, I show how the answers to these three sub questions, together, answer my main research question: *how to ensure responsible innovation of prognostication in postanoxic coma*.

In the following two sections, I explain in detail *how* I aim to answer the outlined research questions.

THEORETICAL APPROACH

The research questions driving this investigation demand two layers of inquiry. At the most straightforward level, the questions are about the medical research practice itself. The main research question reflects this: *'how to ensure responsible innovation of prognostication in post-anoxic coma'*. Yet, at the meta-level, my inquiry requires a methodological reflection on *'how to assess 'responsible innovation' in a clinical setting'*. Drawing from many different fields, this section provides the theoretical background needed to understand the development of the methods I use to answer both questions.

Responsible Innovation

The basic tenet of Responsible Research and Innovation, or Responsible Innovation (RI) for short, is that new technologies warrant societal and ethical deliberation and that these deliberations should then influence the research and development of said technologies (von Schomberg 2011, 2013; Owen et al. 2012; Stilgoe et al. 2013). Moreover, such deliberation and influencing should take place at an early stage when technologies are only emerging, and hence more malleable (Collingridge 1980). Responsiveness to what is considered acceptable or desirable can be driven by principles (von Schomberg 2011) and universal guides such as the declaration of human rights, or, as it is in this research project, it can be directly informed and guided by stakeholders' input (Stilgoe et al. 2013). This perspective on how to responsibly innovate resulted in two increasingly common approaches. On the one hand, anticipatory approaches like scenario-building to prepare for the unpredictable are thought to allow reflexive intervention (Barben et al. 2008, Selin 2008, Boenink et al. 2010, Fisher et al. 2012). On the other hand, the need to investigate actual practices has gained more popularity in recent years. This trend includes the empirical investigation of existing practices, even before consulting stakeholders (Boenink et al. 2016), as well as fieldwork in research practices and laboratories (Fischer 2007, van der Burg & Swierstra 2013). The approach I propose for responsible innovation in clinical contexts in chapter one, lies in between the two. While anticipatory methods risk being too speculative, existing practices are often insufficiently forward looking. Yet, as explained below and in chapter 1, I argue that clinical practices where the research-practice distinction is blurred offer an excellent alternative.

Science and Technology Studies & Philosophy of Technology

My focus on technology development in practices quickly resulted in a fascination for social constructivism and what different approaches actually *do*. Similarly, as put forward in science and technology studies (STS) (Latour 1996, Felt et al. 2016) and philosophy of technology (Verbeek 2005, 2006), technology is not an independent, neutral device that only performs the activity it was designed to perform. Instead,

it should be considered as part of a specific socio-technical network, where a combination of actors, material devices, and the environment they function in, mutually influence one another (Law & Singleton 2000). Depending on the socio-technical network and the way that it impacts the technology, the innovation's performativity, or the way that it in turn impacts the practice, can change. Adopting this conception of technology as part of and interacting with a socio-technical network, I focus on studying cEEG's performative effects and the potential context-dependent variations in its performativity.

A practice-based approach is thus necessary to investigate these two research objects: a) the ways in which technology interacts with the socio-technical network it becomes part of and b) the valuations that emerge in such evolving practices. When investigating practices, I thus aim to identify stakeholder concerns in light of the technology's performativity in a particular context as the basis for RI agenda setting. At the same time, I also aim to deepen these empirical findings by conceptual analysis of stakeholders' concerns, the technology's performativity, and the interplay between the two. As I explain further down, alternating between the practical and the theoretical offers a binocular view on complex issues that I have found enriching.

Liminal Innovation Practices & Comparative Perspective Component

In the method I developed, I thus adopt a practice-based, empirically informed approach of RI. Combining notions from STS and philosophy of technology, I aim for thick descriptions of valuing processes in practice and identifying performative effects of the technology. To get the most appropriate point of reference, I choose to investigate specific clinical practices where the innovative technology has been implemented in the clinical setting as part of research. In chapter one, I call these liminal innovation practices and justify why they are the ideal study object for RI in clinical research settings.

Exploratory research of liminal innovation practices, one in the United States and two in the Netherlands, at the beginning of this project indicated that comparative, multi-contextual research may furthermore be essential to responsible innovation in clinical settings. Medical technologies travel quickly across cultural borders and they express themselves differently in different contexts. Therefore, I set out to investigate the values embedded in current practices of care for postanoxic coma patients in the Netherlands, the USA, and Israel. However, rather than presenting a systematic comparison of the practices in the different countries, I used the fieldwork in the US and Israel to contextualize the findings pertaining to the Dutch practices. In the Netherlands, I examine the role of emerging socio-technical architecture and the possible impacts it may have on care practice, as well as the views of future users (medical professionals, ex-patients, families). In addition, I engage these stakeholders in deliberation on the desirability of such an innovation.

Empirically informed philosophy

While drawing from different fields and disciplines, there is an overarching rationale directing my overall approach. On the one hand, this research project is fundamentally a practice-based, empirical study. The practices studied and the actors involved in these practices are essential to the kind of research presented in this dissertation. Nevertheless, my aim is to combine this ethnographic and qualitative approach with a fundamentally theoretical philosophical approach. The latter is essential for conceptual analysis and elucidating principles and distinctions, as explained in the method section below. Both approaches are important and deserve great merit for their independent value and contributions. Nevertheless, for a project of practical ethics such as this one, using just one of them would be a missed opportunity.

Scholars in theoretical philosophy are often criticized by the more empirical scholars for being 'out of touch' and failing to do justice to the complexity of practical challenges by practicing 'armchair philosophy'. On the other hand, being knee-deep in fieldwork, attempting to make sense of some very complex but narrow issues can easily result in work that is so context-specific that its insights lose value for broader use. Overcoming the chasm that lies in between the two approaches is not self-evident, however, great advantages can result from trying to bridge the gap. The abstraction of (analytic) philosophy often invites artificially removing things from their context (or simplifying their context) so as to better evaluate them, testing them in essence without contextual distraction. However, it is important to test the resulting analysis in its context, once it is put back into place. That's why pushing back on the theory with empirical detail often improves the theory substantially and makes it more robust. Empirical investigation, on the other hand, while having the advantage of seeing things 'in the wild', contextualized with meaning, can allow for philosophers to observe the subject in its complexity. Often, however, this leads to a perspective and understanding that is original in its specifics but that may not be quite as novel, generally speaking, as the investigator might imagine. It is thus important to reconnect that specificity with the hundreds of years of philosophical thought that came before, situating the findings within the larger philosophical debate.

In short, while the aims of RI worked as a starting point, the empirical approach I adopted combines notions of STS and philosophy of technology, as well as my own perspectives on empirical assessment of clinical research practices. This empirical approach is then combined with a more theoretical, philosophical analysis.

METHODOLOGY

In what follows, I give an overview of the methods used. Due to the methodological variety, most chapters have their own method section while the background and motivation for the overarching approach is explained above and in chapter one.

Literature study

As will become apparent from the references used throughout this thesis, I reviewed medico-technical literature in order to understand current prognostic protocols, existing practices, and emerging developments. I consulted bioethics literature to acquaint myself with the existing ethical and societal concerns. These sometimes led to further exploration of philosophical debates in the analytic tradition. Additionally, in order to assess which methods are most appropriate to this context and to fine-tune my own methodology (as stated above), I explored the literature of RI, STS, and philosophy of technology.

Medical ethnography: observations and interviews

Initially, conducting exploratory, ethnographic observations in the United States and the Netherlands helped me identify the kind of practices that would be most beneficial for this research, namely, those that had implemented cEEG for research purposes. I then continued ethnographic, qualitative fieldwork in two such practices in the Netherlands, combining observations and interviewing to investigate both the hidden and implicit concerns as well as the known and explicit ones. This kind of qualitative research is meant to enrich the data and increase the likelihood of exhaustive data collection (Goodson & Vassar 2011). The ethnographic study I conducted in the Netherlands contains a total of twenty patient cases.

My *observations* focus on individual patient cases, following their trajectory from the moment they arrive at the Intensive Care Unit (ICU) until they are either discharged or until they pass away while in the ICU. These processes include the patients themselves, their families, the medical professionals and sometimes a social worker. When observing medical professionals, practitioners as well as research-practitioners, I apply a complementary technique born from the concept of *midstream modulation* 'through which scientists and engineers, ideally in concert with others, bring societal considerations to bear on their work' (Fisher et al. 2006). Specifically, adopting the interview protocol developed by Fisher as part of *Socio-Technical Integration Research (STIR)*⁶ helps map the decisions made and invite the medical researchers and practitioners to reflect on their decisions and the subsequent consequences (Fisher 2007, Fisher & Guston 2008). I choose to do this because exploratory research quickly revealed that my mere presence as an

⁶ I only use the STIR interview protocol, without adopting all other activities and purposes of midstream modulation.

observer has an inevitable impact on the practices I observe. The STIR interview protocol is a constructive way of handling my presence as a researcher because it gives room to the observed professional to explain their underlying rationale, to give potential alternatives, and to reflect on why they choose one alternative over others. Observations resulted in observation reports and fieldnotes. Four cases were selected for in-depth analysis based on their varying outcome predictions and outcomes (see Table 1 below).

Cases	Prediction based on existing practice	cEEG-based prediction	Actual outcome
Reinhart	Poor	Poor	Died after withdrawal of life support
Willems	Good	Good	Lived without severe neurological damage
Vanderbilt	Uncertain / optimistic	Poor	Lived with severe neurological damage
Felts	Uncertain / pessimistic	Poor	Lived without severe neurological damage

Table 1 – Predicted and actual outcome of the 4 case studies (*patient names were changed)

For these four selected cases, I conducted semi-structured, open, qualitative *interviews* with medical professionals, family of patients, and, whenever possible, the patients themselves. Interviews were recorded, transcribed, and translated from Dutch to English. Informed consent was obtained from all individual participants included in the study. Their personal data were anonymized and their names changed. The goal of the interviews is to elicit the subjective experience of the respondents (Weiss 1995). Aside from using Weiss' general notions on conducting this type of interview, I constructed the interview schedules and questions based on suggestions made by Smith and Osborn (2007) in Interpretative Phenomenological Analysis (IPA). To deepen the scope of my understanding of the personal experience of a single event, I interview different stakeholders about the same case, thus adopting a variety of perspectives. For example, per case I interview different relatives of a patient, when possible the patient themselves, the practicing neurologist and/or intensivist, practicing nurse, and, when applicable, the rehabilitation doctor otherwise known as physiatrist (see Table 2).

The respondents' experiences serve a double function. First, they give a deeper and more intimate insight into the respondents' confrontation with, and their being part of, the studied practice; enriching data that is referred to in the literature as experiential knowledge (e.g. Pols 2014; Caron-Flinterman et al. 2005). Second, it provides a mutual reference point, where the interviews can serve as a background for the observations and vice versa, so that differences and similarities can be made visible and studied.

Cases	Reinhart	Willems	Vanderbilt	Felts
Relation to patient (Respondent)	Partner (Irene)	Partner (Judith)	Partner (Johanna)	Partner (Anke)
	Eldest child (Lukas)	Patient (Gerrit)	Eldest child (Natan)	Patient (Anton)
	Youngest child (Jason)	Neurologist (dr Bos)	Youngest child (Roos)	Only child (Jolien)
	IC Nurse (Alice)	Intensivist (dr Berger)	Intensivist (dr Harfe)	Neurologist (dr Helmers)
	Intensivist (dr Valk)	Physiatrist (dr Raven)	Physiatrist (dr Naas)	Intensivist (dr Bent)

Table 2 – Respondents per case study (names were changed)

Stakeholder-informed innovation: workshop, surveys, and expert panels

Aside from engaging stakeholders through their experience, my supervisor, Marianne Boenink, and I organized a *multi-stakeholder workshop* with technology developers, researchers, healthcare practitioners, patients, and family members. Aiming to put different stakeholders together, we united diversified perspectives, in order to jointly discuss a) their desires, needs, and concerns and b) how these can be translated into design and implementation requirements (Krabbenborg 2013). Here the anonymized information from the case studies serves a third function as they are used to design realistic scenarios that were used during the workshop to spark the imagination of stakeholders.

In collaboration with technology assessment experts Janine van Til and Eline Beens and neurologist-researcher Jeannette Hofmeijer, I also participated in quantitative research on this topic. We designed a *survey* for national inquiry regarding public preferences for prognostication in postanoxic coma; the results of which are discussed in chapter 5. The survey was developed and designed, in part, in response to and informed by the data collected during my qualitative study. With its wider range, the survey aims to generalize and, in combination with similar surveys our team conducted abroad, it aims to contextualize the local findings within a broader context.

Finally, our team organized several *expert panels* for valorization of our findings, meaning that we check how valuable the research and our findings really are to the stakeholders in the field. Once a year, we presented the collected data and insights

to a panel of experts and received feedback and input from a variety of medical practitioners, researchers, and technical experts. For broad inclusion, the panel included developers from the new technology as well as medical professionals from both acute and long-term care. For a number of these meetings, their input was recorded and included in the data set for further reflection and analysis.

Comparative Fieldwork: observations and interviews

Exploratory observations and interviews conducted at a research hospital in the *United States*, helped give direction to the study design and offered a contextualized perspective on the study conducted in the Netherlands. The US hospital practice was similar to the Dutch ones insofar as it had implemented the novel technology as part of the care practice for research purposes and thus qualified as a liminal innovation practice, ideal for RI-assessment. Overlap in concerns indicated pervasive and general issues, whereas differences pointed to variable concerns and potential learning opportunities. The latter, especially, command additional reflection and analysis.

For a very different point of reference, I conducted *observations* in four medical institutions in *Israel* where withdrawal of life-sustaining treatment is prohibited. It quickly became apparent that practices widely differ within the country, depending on the geographical location and its demographics. I thus chose institutions in different parts of the country: one in Tel Aviv (relative to Israeli practice, closest to Dutch practices), two in Jerusalem (furthest removed from Dutch practices), and one in Beersheba (somewhere in the middle of the spectrum). I conducted *interviews* with 10 medical professionals: a cardiologist (head of cardiac ICU); an intensive care nurse, a senior intensivist, a paramedic (Ph.D. and teacher of paramedics), an anesthesiologist (head of surgical ICU), a liberal emergency medicine physician (and paramedic), an orthodox emergency medicine physician, the head of acute geriatrics at a chronic ventilation unit, and a nurse practitioner at the same chronic ventilation unit.

While the study in the United States offered a perspective on the potential variety in liminal innovation practices, the study conducted in Israel put the more fundamental, universal challenges in perspective. These contextualized perspectives helped my understanding of the Dutch practice and the analysis of challenges and how best to resolve them.

Conceptual Analysis

In the theoretical approach section above, I emphasized the value of combining empirical work with more theory-oriented, analytic philosophy. The empirical data collected shows immense complexity and appropriately responding to the concerns raised requires, in the first place, a better understanding of them. The

analysis is thus informed by the fieldwork, documented in observation reports, fieldnotes, and open, semi-structured interviews with medical professionals, surviving patients (when possible), and their family members. Conceptual analysis, aimed at clarifying and elucidating, allowed for such better understanding and, as a consequence, revealed quite a few epistemic challenges. Further conceptual analysis and revisitation of the empirical data helped understand how to best address said epistemic challenges. Chapters 2 and 3 (on uncertainty) and the last two chapters (on self-fulfilling prophecies) are evidence of the success of this process. In addition, comparing practices in widely different contexts contextualizes the conceptual analysis and allows for appropriate generalization. This also means that I have limited normative ambitions. My aim is to identify the existing values and concerns and do right by them by enabling improved decision-making through innovation rather than determining what values and concerns ought to be considered. I do, however, explain how I arrive at conclusions as to what is less or more 'responsible' and why.

CHAPTER OVERVIEW

In the following 7 chapters⁷, I aim to combine the empirical and theoretical approaches outlined above in order to develop an epistemic analysis of innovation of neuroprognostication. In chapter 1, I motivate my approach to RI and defend to what extent it differs from existing approaches and why. Chapters 2 to 4 focus on the practices of neuroprognostication in intensive care. First, in chapter 2, I describe the dilemma due to uncertainty as experienced by the respondents in the different case studies as well as the concerns and questions that emerge in practice. I then, in chapter 3, further investigate the two main questions that require answers in order to improve neuroprognostication: 'What outcome can be achieved?' and 'Will such an outcome be acceptable?'. In chapter 4, I then evaluate to what extent cEEG as a new neuroprognostic tool can provide answers, and whether these answers can alleviate the dilemma sketched in chapter 2, in what way, and for whom. The article included as Chapter 5, which I co-authored with colleagues from the departments of Health Technologies & Services Research and Clinical Neurophysiology, presents the public preferences regarding the two questions and the ways in which the new tool can provide answers. In the final two chapters, I focus on self-fulfilling prophecies. Chapter 6 presents a general descriptive analysis of self-fulfilling prophecies in medicine and their ethical and epistemic implications. Chapter 7 further investigates the issues that self-fulfilling prophecies pose when attempting to provide answers to either question in neuroprognostic practice and research.

⁷ While the chapters build on each other and form a united whole, they are also meant to work as stand-alone articles, some of which have already been published. This means that some overlap between them is inevitable.

I conclude by providing answers to the research questions and discussing the pitfalls and opportunities for responsible neuroprognostic innovation. After discussing the wider relevance of my findings, I end the conclusion with an emphasis on the broader epistemic challenges and what can be done to overcome them.

Chapter 1

Liminal innovation practices: questioning three common assumptions in responsible innovation

ABSTRACT

Although the concept of Responsible Innovation (RI) has been applied to different types of innovations, three common assumptions have remained the same. First, emerging technologies require assessment because of their radical novelty and unpredictability. Second, early assessment is necessary to impact the innovation trajectory. Third, anticipation of unknowns is needed to prepare for the unpredictable. I argue that these assumptions do not hold for *liminal innovation practices* in clinical settings which are *defined by continuous transition on both sides of the threshold between experiment and implementation, between research and practice, and between past and future innovations*. First of all, technologies at the center of liminal innovation practices have different characteristics than those typically attributed to emerging technologies. Additionally, the innovation trajectory is significantly different allowing continuous assessment and shaping long after implementation. Finally, these differences demand a reorientation in RI approaches for these cases, away from anticipation of the unknown and uncertain, and returning to observation of the known and predictable.

Liminal innovation practices: questioning three common assumptions in responsible innovation

1. INTRODUCTION

Many Responsible Innovation (RI) scholars make the following three basic assumptions. First, emerging technologies require assessment because of their radical novelty and unpredictability. Second, early assessment is necessary to impact the innovation trajectory. Third, anticipation of the unknown is needed to prepare for what cannot be predicted. In this chapter, I aim to broaden the common understanding of RI by identifying a specific way of innovating, especially in clinical settings, for which these assumptions do not hold – but which nonetheless require critical reflection on, and shaping of, new technologies. I argue on the basis of empirical findings, from performing observations and conducting qualitative interviews. After introducing and developing the concept of *liminal' innovation practices* which are *defined by continuous transition on both sides of the threshold between experiment and implementation, between research and practice, and between past and future innovation*, I suggest that these characteristics require a reorientation in the methodologies and approaches offered in the field of RI. Liminal innovation practices demand a move away from anticipation of the unknown and unpredictable, and returning to observation of the known and predictable.

2. CASE STUDY: RI IN PROGNOSIS OF POSTANOXIC COMA

Ninety percent of out-of-hospital cardiac arrests are lethal and most of the people who are successfully reanimated nevertheless remain unconscious (Taccone et al. 2014). They are admitted to the Intensive Care Unit (ICU) in what is called postanoxic coma. Brain injury remains the most common cause of death in these patients, and survivors likely have suffered neurological damage.

Unfortunately, the prognostic tools to predict neurological outcome are limited. Currently, only two indicators for poor prognosis in these patients exist, both focused on the brainstem.² Death or severe disorders of consciousness like

¹ In Section 5, when we are in a better position to reflect on the significance of this type of innovation practices, I will elaborate on the appropriateness of the label 'liminality'.

² Brainstem reflexes are tested through pupillary and corneal reflexes and through a somatosensory-evoked potentials (SSEP) test.

vegetative state or minimally conscious state can be predicted for some patients by testing brainstem reflexes. Yet through these methods, only a fraction of all patients with poor outcome can be identified as such. In addition, these indicators tell us nothing about the likelihood of a good outcome.

Other technologies (e.g. serum biomarkers, other imaging techniques like CT and MRI) have been researched to improve the multimodal prognostic practice, but the use of electroencephalogram (EEG) seems to hold most promise. In the past, EEG-monitoring led to the identification of specific patterns in brain activity (Scollo-Lavizzari & Bassetti 1987). More recent research with continuous EEG (cEEG) gives reason to believe that the development of those patterns over time can provide a robust contributor to the prediction of both good and poor outcomes in this patient group, possibly within the first 24 hours after reanimation (Hofmeijer et al. 2015; Spalletti et al. 2016).

Nevertheless, it is not evident that cEEG as a new technology emerging in this setting is desirable. First, early prognostication may have an impact on withdrawal of care; for patients with a poor prognosis, early prediction could lead to earlier withdrawal of life-sustaining treatment, thus likely hastening the death of the patient. Additionally, the interpretation of cEEG can be translated to quantified measures or combined with artificial intelligence techniques; the quantification of cEEG and the use of algorithms and machine learning can move the prognostic reasoning out of the healthcare professional's sight.

Many ethical issues arise concerning medical futility. On the one hand, the potential for physiological futility lays bare dilemmas regarding reliability, false positives, inter-rater reliability, reversibility, and self-fulfilling prophecies. On the other hand, normative futility calls attention to existing dilemmas regarding religion, value pluralism, human rights, sanctity versus quality of life, reversibility,³ the danger of biases translated into algorithms and machine learning, and differentiation of care and rehabilitation.⁴ These issues are especially complicated when 'poor prognosis' is not clearly and universally defined (Sandroni & Nolan 2015). In addition, there are also many resource allocation concerns at stake. Think of dilemmas regarding healthcare policy, general policies, cost–benefit analyses, and organ donation.

The many concerns that the novel use of cEEG in this setting raises, unveils the dire need to ensure that this innovation is developed responsibly. They motivate the use of RI by eliciting the following questions: Is implementation of this innovation

³ Reversibility can fall under physiological or normative futility, depending on whether or not it is a matter of denial of potential reversibility or dismissal of the value of projected reversibility.

⁴ The distinction between physiological and normative futility and the issues they raise will be extensively discussed and analyzed in the remaining chapters.

desirable? And if so, under which conditions can it be considered responsible? While investigating how to best answer these questions, I became aware of assumptions in the RI discourse that demand reflection on RI's concept and methodologies.

3. THREE ASSUMPTIONS IN RI

3.1. Background

RI as a field of study has recently gained a lot of traction. The concept and framework of RI (von Schomberg 2011, 2013; Stilgoe et al. 2013), its features (Owen et al. 2012) and dimensions (Stilgoe et al. 2013) developed in close connection with the concept of emerging technologies.

The basic tenet of RI is that new technologies warrant societal and ethical deliberation and that these deliberations should then influence the shaping of said technologies. Moreover, such deliberation and shaping should take place at a stage when technologies are only emerging, and hence more malleable. However, the use and meaning of 'emerging technologies' have varied with time and places where RI has been promoted.

Initially, RI really took off around 2001 when large scale national programs (in the United States⁵ and to a lesser extent in Europe) focused on conducting research and development on nanotechnology; technologies thought to exemplify the 'new and emerging' because of the 'potentially high stakes, uncertainty, and possible adverse effects' (Funtowicz & Ravetz 1993; Barben et al. 2008, 979; Guston et al. 2014, 2–3). Other quintessential examples are genetic modification and synthetic biology, information and communication technology, artificial intelligence, and geoengineering (Owen et al. 2012, 752; Guston et al. 2014, 2). Additionally, RI has been linked to the notion of grand challenges, especially in Europe,⁶ which has foregrounded RI as a way to reflect on how to creatively use these emerging technologies for addressing urgent social and environmental problems (Kuhlmann & Rip 2014). Finally, and partially as a result of the European Commission's promotion of RI through its framework for research (The Directorate-General for Research and Innovation of the European Commission 2013) and funding schemes on both European and national level (e.g. Horizon 2020, NWO, The Research Council of

⁵ The US National Nanotechnology Initiative was launched in 2001.

⁶ The European Union focuses on six Grand Challenges: Health, demographic change and wellbeing; Food security, sustainable agriculture, marine and maritime research and the bio-economy; Secure, clean and efficient energy; Smart, green and integrated transport; Climate action, resource efficiency and raw materials; Inclusive, innovative and secure societies.

Norway⁷), RI has been used for assessing concrete emerging technologies in specific areas like biomedicine and military technology.

In this third context, research proposals for RI became increasingly technology-driven, no longer focusing on the typical emerging technology. All new applications of (emerging) technologies (such as different uses of drones, CRISPR, smart urban energy systems, 'killer robots', personalized medicine, etc.⁸) now become the subject of RI. Funded for a large part by NWO, the Dutch fund for scientific research, the project on cEEG described in the introduction fits into that last context.

Although the following three assumptions in RI persist throughout the different ways in which emerging technologies were contextualized, analysis of this case study shows they may be problematic for this third context of more concrete technological innovations.

3.2. Assumption 1: emerging technologies, radically novel and highly unpredictable

The first assumption is that 'innovation' in RI does not only continue to point mainly to technological innovation,⁹ as criticized by Blok and Lemmens (2015). It seems to also imply a radically novel technology that is either emerging in its conceptual stage or in early stage of research and development. In their article, 'What is an emerging technology?', Rotolo, Hicks, and Martin (2015) identify features such as radical novelty, prominent impact (often described as having disruptive effects) and uncertainty, as well as ambiguity. These characteristics recur throughout the RI literature (for example, Stilgoe et al. 2013, Owen et al. 2012, von Schomberg 2011, Schroeder & Ladikas 2015, Wender et al. 2014, Fisher et al. 2012, and van de Poel 2009, 130). Fisher and colleagues (2012) say, for instance, that their paper on responsible healthcare innovation considers 'some of the potentially disruptive uncertainties surrounding such innovations, focusing on societal and ethical issues often associated with them' (857–858). On account of these features, emerging technologies are thought to be capable of changing the status quo. They are then met with a quite appropriate degree of concern, since the unpredictability of their development causes undetermined, open, and highly uncertain futures.

⁷ ec.europa.eu/programmes/horizon2020/en; nwo-mvi.nl; forskingsradet.no/samansvar

⁸ For a comprehensive list of projects funded by the Dutch RI (MVI) program see www.nwomvi.nl/projects

⁹ RI is sometimes promoted as less technology driven than the preceding 'Technology Assessment' and 'Ethical Legal and Social Aspects (ELSA) or Implications (ELSI)' research. Yet, in practice, the focus on technology in RI remains substantial. This restriction to technology is explained by Godin (2009, 21), whereas the extension to 'innovation' represents a shift from social shaping of technology, to the way the interaction of technology and society shapes (technological) innovation (Grunwald, 2014, 25)

3.3. Assumption 2: impacting innovation trajectories early on

One of the central ideas in RI is that developers, researchers, and stakeholders can impact the development of 'emerging technologies'. This notion is not original to RI specifically, it was adopted from previous strands of innovation studies and technology assessment. The much referred to *Collingridge dilemma* symbolizes this thought well by saying that at the beginning of the innovation process changes can be made easily, but it is hard to know what changes are needed; while towards the end of the process – when we know what changes are needed – it is much harder to implement them (Collingridge 1980). Although the empirical validity of the claim that innovation progresses in a clear and linear way has been criticized by RI scholars (van de Poel 2009, 2013; Fisher et al. 2012) most of RI's ambition is to impact the outcome, or at least the trajectory, of the technological innovation and to do so in early stages of development (Grunwald & Achternbosch 2013). For instance, one of the declared four dimensions of RI (Stilgoe et al. 2013, 1572) as well as one of RI's three main features (Owen et al. 2012) is 'responsiveness' with the specific aim to 'influenc[e] the direction of [research and innovation] and associated policy' (751).

These potential impacts and influences, preferably early in the process, are predominantly based on the idea of a process developing progressively in time, in which there are designated moments of decision-making, between the start of the innovation and its ending. A more explicit example of this is the application of stage-gating in RI (Owen et al. 2012, 756) a project managing technique pushing for artificial division of a project in stages that are separated by decision points. This rationale, including the concept of a clear ending and beginning, betrays the survival of a rather linear conception of the innovation process.

3.4. Assumption 3: the need for anticipation

The third and last assumption follows from the previous two, i.e. because emerging technologies are thought to increase uncertainty and have a potential for prominent impact, and because it is believed that we can impact the course of the innovation, anticipation is thought to be a crucial step in the attempts to make innovation more responsible (Fisher et al. 2012; Owen et al. 2012, 755; Stilgoe et al. 2013, 1570; Guston 2014; Nordmann 2014, 87; Wender et al. 2014; Dickel & Schrape 2017). The aim of anticipation is not prediction, however, but rather to be prepared for what cannot be predicted (Barben et al. 2008, Owen et al. 2012, Guston 2013) and to be able to avoid hazards and obtain advantages as well as possible.

As Lucivero, Swierstra, and Boenink (2011) explain 'The rationale for [early assessment] is that it is preferable to anticipate developments when they are still malleable, even if at this stage the future is still unknown and uncertain (129).' One way to acknowledge uncertainty and anticipate potential outcomes is through

future imaginaries – for example, by constructing scenarios in which speculations, estimations, and educated guesses can be integrated. Better understanding the variations in how the technology can perform and how users may change the innovation's purpose can indeed help the discussion regarding desirability and the need for reflexive intervention (Grunwald 2004; Barben et al. 2008; Selin 2008; Boenink et al. 2010; Swierstra & Keulartz 2011).

These three assumptions as such are very intelligible and justifiable. So much so that they have been adopted throughout the different contexts in which RI is used, even though they may not be as suitable for the above-described third context where very specific new developments are assessed through the RI lens. While I stumbled upon discrepancies between the three assumptions above and the realities of the innovation practice I studied, I recognized that this may be an issue with innovation in clinical practice in general and perhaps with other RI projects that assess concrete technologies.

4. INNOVATIONS IN CLINICAL SETTINGS: QUESTIONING THE ASSUMPTIONS

4.1. Questioning assumption 1: technology meets incremental innovation

In contrast with the first assumption, research projects and developments in medical technology are not typically focused on the development of radically novel technology. Rather, the simple addition of another variable – whether spatial, temporal, contextual, or technological – can lead to the emergence of a technological innovation. cEEG, for instance, consists of a new application of an existing technology. Instead of having the EEG record one moment in time or at intervals, it now records continuously for 72 hours. Other examples of 'new application, old technology' can occur by combining two existing technologies. For example, combining a positron emission tomography (PET) scanner with a computed tomography (CT) scanner led to the creation of the PET-CT. This development, in turn, helped spark the combination of PET (functional imaging) with magnetic resonance imaging (MRI, soft tissue morphological imaging), from which the PET-MRI emerged.

Another common innovation practice in this setting is to adopt an existing application of an already established technology for a new patient group, another illness, or another organ. cEEG, for instance, has recently also been researched as a tool to diagnose delirium (van der Kooi et al. 2012; Naeije et al. 2014). Other examples are the innovative uses of MRI and endoscopes (i.e. flexible tubes with a light and camera meant for internal visual examination). An MRI was first used to detect cancers, but thanks to its ability to differentiate so clearly between grey and white matter, it was further developed for studying the central nervous system and

detect conditions like dementia, cerebrovascular disease, epilepsy, etc. Similarly, there are currently over 10 applications of endoscopes, each of these focusing on a different part of the body.¹⁰

Although these are all examples of visualization and prognostic technologies, the same can be said about more invasive technologies or even some developments in medication. Pacemakers, for example, opened the pathway for implantable cardioverter defibrillators (ICD) and deep brain stimulation (sometimes even referred to as a brain pacemaker). Some medication has side effects that cause off-label use which in turn sparks research on the development of a new drug.

None of these technological innovations represent the kind of radical novelty that is generally referred to in the RI literature. Innovation in clinical settings is more about 'making things new', rather than 'making new things'. And even though these innovations can nevertheless result in prominent impact (by hastening withdrawal of life-sustaining treatment, for instance), they are mostly incremental as opposed to suddenly disruptive developments. Uncertainty remains but to a different degree. The incremental innovations are not necessarily unpredictable in and of themselves. What is feasible in each innovative step is for the largest part already known. It is known, for instance, that EEG is capable of recording brain wave activity. What is still unknown about this technology is its prognostic value in postanoxic coma and the effects it may have. The future is not as open, as undetermined, and, therefore, not as highly uncertain as it is in light of the standard conception of emerging technologies.

A question then worth raising is whether perhaps these examples of technological innovation should not be referred to as 'emerging technologies' at all. The term may have seemed apt for concrete innovative technologies, initially, to justify the need for an RI approach and attract funding accordingly. Since these innovations indeed require an RI lens, identifying them as emerging technologies may have broadened the scope of what 'new and emerging technologies' entails. Their characteristics may have evolved under these funding programs to include other types of innovations. If so, we are left to accept they fall under the term 'emerging technology' despite the fact that they may not share the same traits. If we find this unsatisfactory, we must find different labels for different types of innovations that necessitate RI approaches. The latter being my preferred conclusion, I offer a new term for the innovation I study. However, whether or not we need differentiation in terminology, either way, the kinds of technological innovation that require RI assessment have certainly evolved and it is time RI evolves with them. Therefore, RI researchers must ask themselves time and again whether the various innovations they study actually resemble the ones RI used to focus on in the past.

¹⁰ They range from bronchoscopy for the airways and colonoscopy for the colon and large intestines, to arthroscopy for joints.

4.2. Questioning assumption 2: a continuous lemniscate of innovation

Innovation patterns in clinical settings are certainly not linear. Incremental innovation moves forward, backward, but also up and down. The three-dimensional process can split several ways and is constantly revisited. One can see the development of new experimental applications building on previous ones before the latter are firmly implemented in clinical care. A secondary or incidental finding may spark new research, even when the original research gets abandoned. The transitional phase of innovation, therefore, seems to never end, even if it is sometimes interrupted or side-tracked. The upside is that room for change and impact always remains.

Unlike what the Collingridge dilemma claims, these practices continue to be flexible; innovations remain malleable, even after implementation.¹¹ For instance, since 2002, following two pivotal trials, postanoxic patients arriving at the ICU are almost always treated with therapeutic hypothermia for the first 24 hours in order to prevent further neurological damage (Bernard et al. 2002; Holzer et al. 2002). In many hospitals, this means the patients are cooled down to 33°C. Over a decade later, in 2013, a new international trial showed that 'hypothermia at a targeted temperature of 33°C did not confer a benefit as compared with a targeted temperature of 36°C' (Nielsen et al. 2013)

In one of the hospitals where cooling is now limited to 36°C simply to prevent fever, an intensivist demonstrates during an interview how he explains medical research to families:

Right now, the patient is treated at 36 degrees. Five years ago, I was sitting here in this very room with your predecessors asking them 'would you be okay with your father participating in a study in which we randomize between 33 and 36 degrees' and in five years from now I'll be asking people something else but then the previous thing will have become part of structural treatment. This is how it works. It's a continuous process and a continuous attempt to improve, in which we also implement things every now and then that are worse. Then you find out years later that it has to be different.¹²

About cEEG this respondent then added:

I don't see this as a fixated new development or anything. I always tell family, 'the way your [family member] is treated now, is something we

¹¹ Since my research focussed on practices in research, university, and teaching hospitals, it may be that other hospitals are less flexible. Yet the continuously adapting trend is generally observable throughout national medical practice.

¹² All interviews quoted in this chapter were conducted as part of a larger qualitative empirical study. For present purposes, however, quoted excerpts are not intended to provide substantive support, but rather provide illustration in the voice of experiential experts.

learned 5 years ago'. And aside from the fact that we now apply some EEG things, we're also learning new things and those we'll only implement in two or three years. So, this is a kind of step in between. Perhaps we'll start, well I don't know what other things we could measure, maybe we'll insert a catheter in the retina or whatever, so yeah, this is just an intermediate step.

This respondent voices my concern that one cannot assume a linear trajectory because one cannot assume a clear separation between the experimental phase and the phase of implementation. In part, this is due to the lack of separation between research and care in clinical settings, an issue I will discuss in depth later.

Since clinical care is in constant movement and incremental innovations are found and tested continuously, it is in no way obvious where the innovation process begins or ends. On the contrary, the clinical setting offers fertile ground for a never-ending transitional phase of innovation practices. Even when one specific, incremental step in the process can have a distinct beginning and ending, it can most likely only be identified in retrospect.

For RI, this means that although clear decisions on how to impact the trajectory can be made, decisions can also be withdrawn, adjusted, or adapted. Sometimes impacting the innovation trajectory means moving backward, but more often than not, it involves getting side-tracked. The overarching, organic decision-making process underlying continuous, accumulating innovation is more like a three-dimensional, mostly expanding, sometimes shrinking, lemniscate-shaped¹³ explosion of clusters, rather than a straight line.

4.3. Questioning assumption 3: the need for anticipation in clinical settings

As previously explained, the assumed need for anticipation in RI rests on the first two assumptions regarding the characteristics of emerging technologies and the decision-making processes aiming to impact the innovation's trajectory. If these assumptions are not fulfilled in the kind of innovation described here, should we then not reconsider the need for anticipation in these cases?

Although some characteristics of emerging technologies such as radical novelty and unpredictability leading to high levels of uncertainty may not be representative for technological innovations such as cEEG, the latter's implementation can nevertheless cause prominent impact. Therefore, assessment of the technology, its effect, and desirability is still called for. But since the non-linear course of the innovation allows for more flexible decision-making processes, including

¹³ The lemniscate or infinity-symbol '∞' here symbolizes the never-ending transitional phase of innovation.

the possibility to continuously steer and shape, we may have better ways to be responsive.

One of the advantages here is that the innovation process, including the 'unknown factors', can be observed in real-time, as they emerge. Because research takes place – and can only take place – within the planned context of implementation, we can study the impacts of the innovation 'in vivo'. The use of cEEG for detecting brain damage after cardiac arrest, for example, can only be researched by applying it to actual coma patients in the ICU. The effects of implementation are already expressing themselves, and the assessment of potential pathways can be done based on observations of the care practices during research. Ongoing monitoring and adjusting (Guston & Sarewitz 2002, 99–100; Grunwald 2004, 36; van de Poel 2009, 139–140; Nordmann 2014, 95), even after implementation, are therefore excellent RI approaches.

To be fair, anticipation is not only argued for because of uncertainty about the performativity of the emerging technology. Uncertainty may exist also about how the emerging technology will be used. While social uncertainty in these practices is equally relevant and important to study as it is with the more radically novel technologies, it is again different in degree. In asking 'how will users interact with the technology', we must recognize that the use of the innovation within clinical settings is mostly restricted and controlled. For instance, the possibility for trained healthcare professionals to use the innovative application in a different way than prescribed, will only be permitted (but within that restriction often encouraged) as part of new research, which may then lead to the next incremental step in the ongoing, continuous innovation process.

Finally, I would like to address the sensible counterargument that anticipation as traditionally referred to in RI is nevertheless useful when assessing the overarching trajectory of this type of innovation practices. Even though these incremental innovations are significantly different, in combination or through accumulation, they may nevertheless end up being uncontrollable or unpredictable. The more radical impacts and the unforeseeable, therefore, lie in the overarching trajectory rather than in single innovations.

Not only is this a legitimate response, it is especially important to consider the overarching trajectory in the clinical setting. The controlled, systematic, and incremental innovation processes in clinical settings may, in fact, give the impression that there is no reason to anticipate prominent effects that can come about through accumulation, or that the long-term combined effects cannot be disruptive when we gradually ease into significant change. The application of cEEG may be easier to observe, assess, and shape accordingly, but when deep machine

learning is used in combination with cEEG in order to quantify the interpretation of the pattern development, issues that were considered manageable and appropriate may resurface in ways that are no longer acceptable.

One just needs to remain aware that the ultimate outcome of all accumulated increments of innovation in clinical settings (some of which are happening simultaneously) is much harder to observe or control than its individual parts. In addition, the overall outcome may be much more disruptive and unpredictable. These overarching trajectories are fertile ground for the technologically induced change of the moral landscape, otherwise known as techno-moral change (Swierstra et al. 2009), and, as such, they need to be studied. Although observing any single emerging technology in order to discover immediate pathways is equally necessary, it tends to keep the dangers of more distant or combined effects out of sight. However, this is not a matter of *either-or* but one of *both-and*. We need to do both, and ask ourselves *what* is needed *when*? However interesting and useful scenario-building approaches, for instance, may be for the overarching trajectory, they are not very helpful for the incremental, singular innovation. Clearly, then, my criticism of the typical modes of anticipation only applies to the RI assessment of the latter.

5. BETWIXT AND BETWEEN: LIMINAL INNOVATION PRACTICES IN CLINICAL SETTINGS

Thus far, I demonstrated how innovations in postanoxic care and other clinical innovations like it are different from innovation processes usually studied under RI and how those differences challenge three common assumptions. As I previously suggested, a new term for this type of innovation would offer conceptual clarity.

While studying innovation in prognosis of postanoxic coma, I identified some distinct characteristics of what I will call liminality, which I believe are the reason why this innovation practice does not validate the assumptions described above.

'Liminality' is a rather uncommon term derived from the Latin word 'limen', meaning threshold. The term 'liminal' first appeared in 1884, in a psychology publication (La Shure 2005). In 1909, Van Gennep introduced the term to the field of anthropology through his work 'Les Rites de Passage'. However, it was Victor Turner, another anthropologist, who really publicized the term through his theory of liminality. As Turner states,

Liminality may perhaps be regarded as the Nay to all positive structural assertions, but as in some sense the source of them all, and, more than that, as a realm of pure possibility whence novel configurations of ideas and relations may arise. (1967, 97)

Some dictionary definitions of liminal¹⁴ are (1) occupying a position at, or on both sides of, a boundary or threshold; (2) between, or belonging to two different places; and (3) relating to a transitional or initial stage of a process. Similar to Turner's account of the status of liminal individuals, I too argue that the status of liminal innovation practices in clinical settings is socially and structurally ambiguous.

5.1. Between research and care, between experimentation and implementation

First, liminal innovation practices are placed at, and on both sides, of the threshold between research and care. Researchers and policy makers may assume or even want a distinctive separation, but the boundaries between the two are often much more complicated (United States 1978; Beauchamp & Saghai 2012). For innovation practices in the clinical setting, the research-practice distinction is often an artificial one. Ideally, the researched technology or application is not at all present in care. Here, however, there is no other way than to apply the researched technology to the patient within the care practice¹⁵ which is also the anticipated context of use. In other words, partial implementation is a condition for experimentation.

Even so, the idea remains that the technology is only present for research purposes. This ideal is often pursued too. In one hospital's ICU, for instance, EEG machines were connected and put in the patient's room, but facing the wall, preferably with the monitors turned off, and covered with a sheet so as not to interfere with standard care. The cEEG monitoring was thus kept out of sight, quite literally. Nevertheless, after research resulted in several publications, before new clinical guidelines were drafted, healthcare professionals told me of their interpretation of cEEG monitoring and the basis for that interpretation. One intensivist, after being asked whether it was helpful in any way to have the EEG continuously monitoring, responded:

Yes, sure. It's exciting! I peek every now and then. If it starts showing something nice again after 12 hours, or if it's still dramatic stuff after 24 hours, then those are things that make me say, in my opinion it's all going to work out or, it's not going well at all. [...] But I'm no neurologist, I can't really interpret those EEG-patterns, I can more or less,...I've got a bit of an idea from all the patients that lie here, of course; the images that come passing by. [...] But okay, perhaps I should take a course sometime.

Another factor adding to the liminal character of the innovation practice is that researchers often remain part of the regular clinical practice alongside their research

¹⁴ The dictionaries I consulted are: Oxford dictionaries and the Cambridge English dictionary. I included all results.

¹⁵ One can imagine that, in the future, alternative ways – such as computer modeling, virtual reality techniques, etc. – would be possible, rendering innovation through a clinical setting potentially irresponsible.

activities. When treating a patient, they ought to not include their knowledge about the EEG-patterns. Yet how does one's professional opinion remain untouched by one's scientific findings when the use of a new technology is entering an intermediate stage, suspended between care and research?

When technology can be researched only by applying it in the anticipated context of use, users find themselves midway between experiment and implementation. Healthcare professionals, developers, and policy makers start thinking along, whether for or against the innovative practice. Questions, comments, concerns, and feedback may not come only from the medical personnel, but also from patients and family members. 'What's this for?' 'If it's there, why don't you include it?' 'It must say something.' One of the family members of a patient gave an extensive description of their interpretation of the cEEG and the effect it had on them:

If I had seen two or three of those thingies just following a straight line or not producing scores continuously, or one producing a lot and the other very little – because I just saw the same kind continuously. They were all... There was not one outlier up or down or a straight part or anything. I obviously stood and looked at it for quite some time. I was thinking, what is this thing telling us? And really, I found that at least some of the brain was still working everywhere. To me that meant nothing but good news. That was my conclusion. [...] That's the idea I had. And this gave me hope. [...] It gave me hope that there was activity everywhere, brain activity. Even if it was just a little bit and even if it wasn't equally high across the board. But it gave me hope.

Even a clear distinction between which aspects serve care and which serve research does not really work in the long run. The two practices are not just intertwined or inseparable; they are intrinsic to one another. There would be no need for researching a new prognostic tool for postanoxic coma, if healthcare professionals and families did not suffer from the uncertainty regarding the patient's cognitive state and its consequences. And only with patients who are actually in postanoxic coma, can the most productive research on prognostication of such a coma take place. Both practices are in constant communication through feedback loops. Consequently, they cannot help but shape one another. And this cyclical pattern brings me to another liminal characteristic.

5.2. Between previous and future innovation

These innovation practices embody a transitional stage, not only between research and care and between an experimental phase and a phase of implementation, but, additionally, between previous innovations and future ones. As described earlier, the constant reinvention of medical technologies and their uses causes a seemingly never-ending transitional phase where one innovation serves as a building block, or a kick starter, for another one. In this process, each separate incremental innovation

may get embedded fully, partially, or not at all. Secondary findings may lead to new innovative research. For example, EEG was originally developed simply to capture electrical activity in the brain. Its use in the clinical field started only after epileptiform spikes were demonstrated as an indication of epileptic activity. EEG had been pursued for the prognosis of neurological damage after cardiac arrest before, which in 1987 only led to the distinction of five rough categories of brain activity (Scollo-Lavizzari & Bassetti 1987). By exploring cEEG recordings, researchers discovered the significance of patterns developing over time. Alongside this evolution, though, other applications of (c)EEG were explored, researched, developed, and implemented or dropped. As one respondent explained: 'I think we'll be able to get more out of the EEG than we already do now. And I think there are new things coming at us that we don't know yet.' The combination of cEEG with deep machine learning is already confirming that idea.

Although each of these innovations represents only an increment of a much larger, overarching evolution, each novel application of EEG is a technological innovation in its own right and needs to be assessed as such.

5.3. Not yet used, yet not 'not used'

Finally, in such innovation practices the researched technology (or application) is not yet used in care, but it is not 'not used' either. As strange as that may sound, this kind of duality where two seemingly opposing views of reality coincide is not entirely uncommon. Observing the 'use' of cEEG during research reminded me of the wave-particle duality concept for explaining the phenomena of light, about which Einstein and Infeld (1938) wrote: 'We are faced with a new kind of difficulty. We have two contradictory pictures of reality; separately neither of them fully explains the phenomena of light, but together they do.' Or, as Turner (1969, 95) says, 'Liminal entities are neither here nor there; they are betwixt and between the positions assigned and arrayed by law, custom, convention, and ceremony'.

The moment a technology or its application exists, it creates new choices that were not there before and, therefore, can no longer be ignored. Once the choice architecture has been changed, even merely ignoring the technology and disregarding its effect is a choice. In these particular innovation practices, the new choices are emerging within the environment where they (will) ultimately matter. The fact that users find themselves in this liminal space causes ambiguous responses to these choices.

For example, one of the patients I followed, under the current protocol would be labeled as a 'grey zone' case, meaning it cannot be predicted whether this person will come out on the good or bad side of the recovery spectrum. The innovative application of cEEG, however, predicted bad neurological outcome. As a result, the

doctor communicated to the family that 'a brain film' was recorded which, following the newest research indicated a negative outcome but since this technology was not yet established, they would give it another week following standard care protocol. Here, the doctor is explaining how he's not 'using' the technology, while nevertheless using it, if only to prepare the family for what is to come.

At some point in the innovation process, the use might ultimately tip over to full integration in the clinical practice, but it is a gradual evolution that does not occur everywhere simultaneously. For example, another intensivist from another hospital said in an interview 'Now it's the case that when the EEG is really very good after 12 hours, I tell the family that "well, it certainly looks good. I'm confident that it will be alright."' When asked if he would inform them also about a negative prediction coming from the EEG he said yes. This kind of communication does not necessarily mean that decisions on whether or not to withdraw care are already, or explicitly, based on the innovative technology. Those statements do, however, actively contribute to cEEG playing an increasingly greater role in the care setting, while moving toward full implementation.

The structural place of cEEG monitoring is ambiguous too. In a hospital where it has been decided not to use cEEG in prognosis (thus solely for research), monitoring seems nevertheless to be part of the standard care protocol. When the wife of a patient asked healthcare professionals what the EEG was for, the response she got was, 'Oh, we always do this, it's kind of standard procedure. We usually do it right away, but earlier there was no machine available.' She did not get an answer about what the machine was doing though, and her impression was that the people she asked did not really know either. I checked with several IC-nurses and many of them indeed were not sure exactly what the EEG's precise function is.

Working in anthropology, Turner explains that liminal individuals have 'no status, insignia, secular clothing, rank, kinship position, nothing to demarcate them structurally from their fellows' (1967, 98) Similarly, cEEG monitoring seems to have no clear status nor can it be clearly differentiated from other applications that are fully embedded, or from those that are purely there for research.

Perhaps the most fascinating and telling factor about this ambiguous status is that cEEG monitoring as a so-called standard procedure complicates the concept of informed consent. In an interview with the previous respondent's husband, he says 'They can research all they want. As long as it is to advance medical science, great! But they could've told my wife and children what they're doing, what kind of research it is.'

When pressing researchers on this matter, they admit that it is a bit of a 'grey zone'. Monitoring is non-invasive, the patient is unconscious, and the family often has

not arrived yet. Since starting the cEEG as early as possible is regarded as part of providing the best possible care for the patient, it is assumed that the family and patient give consent for that, the moment they step into the hospital. The wife of the recovered patient agrees with that premise:

Of course, by going to the hospital you're in fact giving permission to examinations and treatment anyhow. But in that case, one would hope, that when something comes out of that examination, we will be told. [...] Even if these are preliminary findings.

This case reveals the ambiguity between what counts as a research project, and what does not. Even when a research project ends, the monitoring – and subsequent data collection – often continues. Policy for research projects on human subjects is well-defined, but there is no policy for how to deal with innovations that are semi-implemented but no longer systematically researched. Is it then research, or is it standard practice? The questions the respondents raised are legitimate. If it is research, why is there no informed consent? If it is standard testing, why are they not given the results?

I argue that, due to its liminal characteristics, innovation in postanoxic care and especially the use of cEEG monitoring as a way to improve prognosis, is a remarkable example of a *liminal innovation practice—an innovation practice defined by continuous transition on both sides of the threshold between experiment and implementation, between research and practice, and between past and future innovation*. Seeing that other innovation practices in clinical settings have similar features that contradict the aforementioned assumptions in RI, it is likely that many of them are liminal, too. I do not have the empirical evidence to back this claim; nor can I judge whether liminal innovation practices exist outside of clinical settings. Yet unless this case is a peculiar anomaly for reasons I have been unable to notice, I assume other scholars will recognize these liminal characteristics in their own RI research.

6. CONCLUSION: MEANING FOR RI APPROACHES

If the characteristics of emerging technologies justify anticipation of a spectrum of possible futures so that early assessment and shaping of these technologies can happen accordingly, what do liminal characteristics change? What implications do liminal innovation practices have for the methods and approaches in RI? I do not provide a systematic analysis of RI approaches here, but simply offer some general directions that I found especially helpful in my methodological reorientation to assess liminal innovation practices. To do justice to the field of RI, a more thorough account would be required.

First, liminal innovation practices require practice-based RI approaches, away from anticipation of the unforeseeable unknown, back to observation. Similar pleas already exist in RI in general (Fisher et al. 2006, Boenink 2013, Guston 2013, Nordmann 2014, Boenink et al. 2016), and in the case of liminal innovation practices it would be nonsensical to plead for anything else. When research can take place only in clinical settings with actual patients, the effects already express themselves during research, meaning that we can study impacts as they emerge. Moreover, by approaching these innovation practices anywhere outside of the research setting, researchers would miss the opportunity to scrutinize the developments as they unfold. This point is also made by van de Poel with his 'Responsible Experimentation' (2009, 2011), which advocates responsibly testing a technology in society precisely because 'uncertainty and ignorance are partly irreducible and, in as far as they can be reduced, such reduction is an ongoing process' (2013, 353). Similar concerns have also been expressed by real-time technology assessment advocates Guston and Sarewitz (2002). Since liminal innovation practices are often ongoing, continuous monitoring and reviewing is indeed of paramount importance. In light of this, one might think that post-market surveillance (Jacobs et al. 2010, 110; 1; 3) provides a solution. However, it is impossible to apply this ordinarily useful method to liminal innovation practices, because the ongoing overlap between practice and research implies there is no clear and proper market introduction. Precisely because there is no distinct post-market phase, continuous monitoring of liminal practices should be encouraged.

Second, although I endorse RI's focus on stakeholder engagement, the way in which engagement can be organized differs. Liminal innovation practices enable stakeholders to be included through their experience. The importance of experiential knowledge, while sometimes hard to capture, has been increasingly acknowledged and illustrated by researchers in the field (e.g. Pols 2014, Caron-Flinterman et al. 2005). This knowledge is automatically generated in liminal innovation practices. It would be a waste not to capture it. As shown throughout this chapter, qualitative interviews with stakeholders who've been part of, or confronted with, liminal innovation practices reflect in-depth knowledge about how innovations express themselves in their context of implementation. Another practice-based approach of engagement is Fisher's 'midstream modulation' (Fisher et al. 2006). This 'governance from within' aims for participation that sparks self-reflection. Since RI researchers often change the practice they observe simply by being present, it may be more reflexive to choose consciously and responsibly what the impact will be. In my experience, using interview protocols that probe for reflexivity (Fisher 2007) is a constructive way of handling one's presence as a researcher. Additionally, it allows for deeper understanding of the observed professionals by giving them room to explain their underlying rationale, give potential alternatives, and reflect on why they chose one alternative over another.

Third, aside from a reorientation towards the study of practices and practice-based inclusion of stakeholders, I suggest doing multi-contextual fieldwork. What really shows when studying a liminal innovation practice is that the effects of new technologies are significantly different depending on the context in which they appear. Therefore, studying multiple sites will increase our understanding of the spectrum of potential impacts. Additionally, including widely different contexts in our assessment, for instance by doing comparative research in countries with a very different cultural background, can help RI meet the demands of value pluralism. For innovations in clinical settings where the ease of adoption is remarkably high, regardless of the cultural or contextual differences, this is crucial. Modern hospitals all over the world have very similar equipment. When a new application of EEG for example, arises in one part of the world it can quickly be used in another. Fieldwork that takes account of practices in the global South – and the far East for that matter – can help decrease the risk of ideological coercion (Macnaghten et al. 2014, 192–193). Last but not least, a lot can be learned about how the introduction of a new technology can change our values by looking at widely different contexts and cultures. Studying a variety of settings where different values apply can be another way to explore and thus anticipate technomoral change.

In short, for RI assessment of liminal innovation practices, I suggest practice-based approaches that include continuous monitoring and reviewing, observations, and in depth, qualitative interviewing in multiple, preferably contrasting, locations.

7. DISCUSSION: DIFFERENT CONTEXTS REQUIRE DIFFERENT APPROACHES

The three assumptions I question in this chapter are not generally problematic, but when RI is used in a different context like that of distinctive technological innovations, we must consider whether the assumptions thus far guiding RI still apply. I do not think liminal innovation practices are the only kind of innovation that require a reorientation in RI approaches. It just happened to be the one that sparked the following important questions: Does the kind of innovation I study sufficiently resemble the kind of innovation the RI framework was initially developed for? And if not, what characteristics are different? Which assumptions do not make sense for the case I'm studying? And what does this say about the method I must employ to assess this innovation? In order to fully grasp which consequences liminal innovations should have for selecting and adjusting RI methods, further exploration is required.

With this chapter, however, I hope I have shown that different contexts may demand different approaches. It is, therefore, crucial to reassess our methods in light of the particularities of the contexts in which RI is used. Anticipation of

potential futures may be useful for emerging technologies like nanotech, artificial intelligence, synthetic biology, or geoengineering. But the need for anticipation of unknown futures has yet to be tested in the face of other innovative practices like liminal ones. This may be true for other assumptions we have come to take for granted as well. Due to the variety of technologies studied in RI, I suggest adopting a critical approach to any future RI-project and the method(s) it requires. Such an attitude will further encourage a flexible concept of RI that allows reinterpretation when new contexts demand it.

Chapter 2

Braving Uncertainty After Cardiac Arrest:

how two questions without definite answers shape a clinical dilemma

ABSTRACT

When patients in coma after cardiac arrest are predicted to have a poor neurological outcome, life-sustaining treatment is commonly withdrawn in response, with the patient's death as a result. Neuroprognostication, however, is precarious. Due to neuroprognostic uncertainty, a moral dilemma emerges: *Withdrawal risks that a patient with good neurological prospects dies, while continuing life-sustaining treatment risks that a patient with unacceptably poor neurological prospects survives.* When deciding between withdrawal, continuation, or advancing/postponing the decision, each choice requires careful consideration of potential harms to the patient, the patient's family, and society as a whole. However, my empirical investigation shows that all of these considerations depend not just on the reliability of the predicted outcome but on whether that predicted outcome correctly projects the harms and benefits at stake. Analysis of the decision-makers' concerns and how they weigh them show that the avoidance of potential harm for a particular patient necessitates answers to two very different questions: *what outcome can be achieved* and *will that outcome be acceptable?* Efforts to reduce neuroprognostic uncertainty, however, focus solely on answering the first question. Further investigation is required regarding what outcomes are found acceptable, how, and by whom.

Braving Uncertainty After Cardiac Arrest:

how two questions without definite answers shape a clinical dilemma

1. INTRODUCTION: AIM AND MOTIVATION

Mr. Reinhart¹ has been without oxygen for over twenty minutes when, following successful reanimation, his heart starts beating again. As most resuscitated patients, he nevertheless remains unconscious and is brought to the intensive care unit (ICU) in postanoxic coma, meaning coma after oxygen deprivation. There, his wife and two sons wait in anticipation to find out whether or not he will awaken in the days to come, after the cooling procedure and sedation are stopped. All three later say that they were aware that their loved one had likely suffered brain damage during the cardiac arrest. For about half postanoxic patients this brain damage proves lethal (Hasager et al. 2018, 989). Mr. Reinhart dies after one week, due to withdrawal of life-sustaining treatment, following a predicted poor neurological outcome which was considered undesirable by medical standards.

This real-life scenario is common. As mentioned in the introduction, withdrawal of life-sustaining treatment (WLST) is, in fact, the major in-hospital cause of death of resuscitated patients after the first twenty-four hours following the arrest (Hasager et al. 2018, 989, 994; Dragancea et al. 2017). The worldwide trend of legalizing WLST for these patients has moved deliberations about the acceptability of WLST from the courtroom to the ICU. As a result, the responsibility for decision-making regarding WLST has now shifted to the healthcare professional, the patient's next of kin, or both. They now have to decide whether or not to withdraw life-sustaining care from this particular individual.

In the Netherlands, following the Medical Treatment Contract Act (WGBO), medical treatment should only be continued if it is in the best interest of the patient *and* the patient agrees to it (Dutch Ministry of Health 2020). For patients in coma at the ICU, this means that the medical professionals are responsible for decisions on WLST (Royal Dutch Medical Association 1997, 2005; Gerritsen et al. 2009). It is, however, common practice to seek information about the patient's preferences through family members and while it is desirable and encouraged to get the family's approval for WLST, their permission is not required (Royal Dutch Medical Association 1997, 2005; Gerritsen et al. 2009; Dutch Ministry of Health 2020).

¹ Mr. Reinhart was one of the patients I observed after his admission to the ICU. All patient names have been changed and personal data have been anonymized.

Whether made jointly or by clinicians alone, decisions are necessarily informed by whatever neuroprognostic information is available to predict what outcome can be achieved when maximized care is continued. However, neuroprognostication for postanoxic patients in coma is precarious and, in practice, the decision-making process is far from straightforward. Innovative prognostic technologies are developed with the expectation to reduce uncertainty and better inform decisions in the best interest of the patient. Yet, with the risk of error looming, the question remains 'how decision-makers in practice decide whether to continue or withdraw life-sustaining treatment when facing varying degrees of uncertainty?'.

In this chapter, I analyze how decision-makers in Dutch intensive care units experience the uncertainty regarding patient outcome in clinical practice for postanoxic patients after cardiac arrest and, subsequently, how they weigh concerns in light of that uncertainty. After describing the methodology of the empirical study (§2), I first demonstrate the general prognostic uncertainty experienced by both the family of the patient and the medical practitioners (§3). Second, I present the dilemma medical professionals and relatives face and lay out the difficulties each choice entails (§4). I then offer a reconstruction of the different considerations and concerns (regarding the patient, caregivers and family, and resources) that are taken into account, all of which point to a desire to avoid the greatest potential harm (§5). Finally, I conclude that accurately predicting harm requires answers to two different questions (§6).

2. METHODOLOGY

I conducted a qualitative, ethnographic study of neuroprognostic practices for postanoxic coma patients in two Dutch teaching hospitals that had introduced a new neuroprognostic tool in their ICUs for research purposes. While the new tool was not yet used to inform decisions on WLST, the experimental evidence was still sometimes communicated to prepare the family, especially in cases where no other information was available. In the previous chapter, I described such practices with blurred research-practice distinctions as *liminal innovation practices*. Investigating liminal innovation practices allowed me to investigate decision-making processes in contexts where a) decision-makers are confronted with the amount of uncertainty experienced in existing, standard practice while b) the potential impacts of changing levels of certainty following the introduction of a new neuroprognostic tool are simultaneously revealed.

The purpose of the fieldwork was to get a better understanding of the existing and emerging neuroprognostic processes and the concerns these elicit in practice. To this end, I conducted ethnographic observations as well as qualitative

interviews. For the analysis, I combined inductive and deductive approaches, starting inductively with the general question: ‘what are people concerned about?’, allowing matters of concern to emerge from the data. Once I developed a list of concerns, I analyzed mostly deductively to further analyse the concerns. The majority of concerns in neuroprognostic research and care that emerged from the data revolved around one central theme: *uncertainty*. In this chapter, I focus on the way uncertainty is experienced in practice and what medical professionals and patients’ families take into account when making decisions under varying degrees of uncertainty.

As I explained in the introduction, I selected four cases for in-depth analysis:

Cases	Prediction based on existing practice	cEEG-based prediction	Actual outcome
Reinhart	Poor	Poor	Died after withdrawal of life support
Willems	Good	Good	Lived without severe neurological damage
Vanderbilt	Uncertain / optimistic	Poor	Lived with severe neurological damage
Felts	Uncertain / pessimistic	Poor	Lived without severe neurological damage

Table 2.1 – Predicted and actual outcome of the 4 case studies (*patient names were changed)

As indicated, the four case studies above were marred by varying degrees of uncertainty. For two cases, a prognosis based on existing practice was available. For the other two, there was not. For those that remained highly uncertain, continuous electroencephalogram (cEEG) monitoring, an innovative neuroprognostic tool under investigation, predicted poor outcome. While this experimental prognostic information was communicated to the family to prepare them for an unfortunate outcome, care was continued following existing protocol. One family remained optimistic rejecting the poor outcome prediction, while the other accepted the poor prognosis but still remained hopeful, despite the pessimistic outlook.

In total, I interviewed twenty respondents in the Netherlands, five stakeholders per case (see table 2.2). When possible, I interviewed the patient, their offspring, and always their partner. Several medical professionals were interviewed for each case.

Finally, the fieldwork I conducted in the United States and Israel, observing medical practices and interviewing medical professionals there, was used for comparative purposes. This helped contextualize my findings and place them in a wider perspective. Although this chapter focuses on Dutch practice, I mention contrasts and similarities between Dutch practices and those abroad where relevant.

Cases	Reinhart	Willems	Vanderbilt	Felts
Relation to patient (Respondent)	Partner (Irene)	Partner (Judith)	Partner (Johanna)	Partner (Anke)
	Eldest child (Lukas)	Patient (Gerrit)	Eldest child (Natan)	Patient (Anton)
	Youngest child (Jason)	Neurologist (dr Bos)	Youngest child (Roos)	Only child (Jolien)
	IC Nurse (Alice)	Intensivist (dr Berger)	Intensivist (dr Harte)	Neurologist (dr Helmers)
	Intensivist (dr Valk)	Physiatrist (dr Raven)	Physiatrist (dr Naas)	Intensivist (dr Bent)

Table 2.2 – Respondents per case study (names were changed)

3. UNCERTAINTY AND THE STRUGGLE TO PREDICT AND CONTROL

The concerns most frequently expressed by those involved in postanoxic care revolve around the pervasive uncertainty and lack of control that both family members and health care practitioners experience. The themes that recur time and again when people describe the difficulties they encountered in the ICU reflect the affective dimension of *not knowing*: wanting to know what is going to happen, wanting certainty, feeling uncertain because of risk, the constant doubt, no way of knowing for sure, hoping for the best, fear of false hope or expectations that won't be met, fear of making the wrong decision, ... Both clinicians and family members described an experience of posing questions without answers.

Family members clearly voice the anguish that stems from not knowing what the outcome will be and having no control over such an unpredictable situation.

Felts' partner, Anke:

The helplessness and the powerlessness. That you don't know how or what. [...] Can someone even wake up at all, and how will they wake up? Will he come to die? What if he won't wake up and you're faced with the choice, are we going to continue with this or will we end someone's life by ending treatment? Will he wake up and live like a vegetable? Will he be severely disabled? But will he be conscious of that? All those thoughts go through your head.

Her daughter, Jolien:

You want to know, is this a good sign, is this a bad sign? Because you have no grip on anything, you have no idea what to expect or what you're going to face. [It's so important] that someone tells you what is happening, what they're doing, but [...] they don't know either. There is no prediction, which is of course what you really want.

Family indeed struggles to understand what is happening, what the test results mean, especially regarding future prospects and what possible outcomes can be expected. At the same time, physicians often have difficulty explaining why there is so much uncertainty, even from an expert's point of view. Although common in medicine and especially so in intensive care, the process of facing probability rather than certainty, doubt rather than knowledge, is something even professionals have to learn over time. Felts' neurologist, dr. Helmers, says: 'You're always a little uncertain just because we simply don't work with certainties. It's always about probabilities.'

Reinhart's intensivist, dr. Valk, further clarifies:

This job is made up of doubt and so you try to take away the doubt. It could be this, so you make another scan. You try to get more certainty, but you're almost never going to get that. They don't teach you that [in medical school]... this constant doubt... What outsiders can't understand is that every test you do, every result you get, is in light of a probability diagnosis. Sometimes, you can't find it, you don't see it show up on a scan, but that doesn't mean it's not there. That kind of uncertainty, it's really difficult and people don't get it. Someone will say 'but you've done the test'. Yeah, but that doesn't mean I'm going to find it.

Most physicians I spoke with confirm that the presence of constant doubt is something they had to learn to cope with over time. Since it was not taught, it is not something they are equipped to deal with when they are fresh out of medical school. Dr. Helmers elaborates on that learning trajectory: 'Back then I may have still had the hope that, one day, I would be all-knowing. That hope I no longer have. Now I realize "no, this is just how it is and the doubt will always remain."' A big part of the lesson is to accept medicine's limitations.

Willems' intensivist, dr. Berger, makes this explicit:

Our predictive capacities are just really limited. For us, at the ICU, as well as for rehabilitation doctors, neurologists,... it all contains an enormous component of reading the tea leaves... yeah, we're just really bad at it. And of course, that's a very difficult message for the family.

On the one hand, these testimonies show that without certainty, the possibility for deliberate action seems to be undermined. Many respondents feel that, in order to act purposefully, one must at least have some degree of certitude about the effects of the treatment; what range of outcomes it may lead to. Especially with regard to life and death decisions, uncertainty can stall decision-making. In fact, the survey presented in chapter 5 which our team conducted among Dutch citizens, shows that half of the respondents find WLST unacceptable when there is *any* uncertainty regarding long-term outcome. Combined, the testimonies and the survey may explain why the powerlessness experienced by both family and caregivers in the face of uncertainty and their constant struggle for information, clarity, and accuracy are indeed some of the greatest drivers in the search for better predictors of outcome in postanoxic care.

On the other hand, the testimonies as well as the effort in prognostic innovation seem to imply that the kind of uncertainty that needs to be addressed is singular and generic: uncertainty about what outcome to expect. Decision-makers want to know what outcome to expect so that, based on whether or not that outcome is desirable, they can decide to either continue or withdraw life-sustaining treatment. However, as we shall see in what follows, this generic notion of prognostic uncertainty is flawed. While it is often difficult to determine what outcome to expect, it is often just as difficult to determine what outcome should be pursued.

4. MAKING DECISIONS IN THE FACE OF UNCERTAINTY: THE DILEMMA

The fieldwork and testimonies revealed that when a patient's neurological outcome is uncertain, both the decision to withdraw and the decision to continue life-sustaining treatment hold great risk. Decision-makers are forced to weigh two potential harms:

Either: withdrawing life-sustaining treatment with the risk that a patient who could have at some point regained consciousness with acceptable quality of life, will now die.

Or: continuing care, allowing the patient to survive, perhaps in ways that they would not have consented to and thus causing them additional suffering.

Medical professionals and next of kin must then decide whether to err on the side of withdrawal or err on the side of continuing life-sustaining treatment.

In what follows, I use the different predicted and actual patient outcomes in our cases to illustrate how the options are experienced in practice. In addition to withdrawing (§4.1) or continuing life-sustaining treatment (§4.2), some contexts will allow, or even force, postponement or advancement of the decision-making (§4.3). With each choice, decision-makers face difficulties, risk of harm, and potential second-guessing.

4.1. Choosing to withdraw treatment

Felts' neurologist, dr. Helmers:

Of course, you're always slightly uncertain when you're going to tell someone 'it's looking really bad'. Then you're saying something that has far-reaching consequences [...] Especially when the moment comes that you're attaching decision-making regarding the cessation of life-prolonging treatment to it [...] Ultimately, you're then deciding over life and death.

Deciding to withdraw life-sustaining treatment based on a poor outcome projection, typically leads to only one scenario: the death of the patient. This makes a poor outcome inevitable, regardless of whether the patient would have died or survived otherwise. The implications of such *self-fulfilling prophecies* require extensive attention and are discussed in the final two chapters (see also Mertens et al. 2021).

In practice, I observed two motivations for deciding to withdraw, once there is sufficient reason to believe that a patient's chances are bad. First, if the patient's death is considered unavoidable, there is no need to prolong the suffering of the patient, their family, or both. Second, if the patient risks surviving in ways that the medical professionals or relatives think the patient would consider 'worse than death', discontinuing treatment seems the preferable care protocol, best reflecting the patient's interests. Note how the qualitative description of outcomes 'worse than death' points not to what outcome can be achieved but the way that outcome could possibly be evaluated. The metaphysical character of 'the will of the unconscious patient' does not make the identification of treatment goals any easier either.

Vanderbilt's intensivist, dr. Harte, elaborates on the treatment goals of intensive care, one of the core pressing concerns in neuroprognostication:

What matters to us, is that people who receive our treatment at the ICU are getting a life back that they can move on with. It may not be the life they were used to, but a *life* they can go on living. That is what's important to us. Maybe some things won't go well anymore or they'll struggle with some things, that's possible. But vegetables, that is not what we deliver. That's not what we strive for.

Medical practitioners are essentially trained to provide their patients with care that leaves them better off than if they had not received said care. Escalating care when it risks harming the patient feels antithetical to the professionals I spoke with. Although some, especially in the United States and Israel, resort to terms like futile care, most Dutch practitioners describe care past a certain point as *nonsensical*. The greatest struggle for them is to know where to draw the line. 'Where's the limit-- that's a difficult question', says Reinhart's intensivist. With this difficult question, the

intensivist already points to the uncertainty revolving around the acceptability of an outcome, regardless of whether the eventual outcome can be determined.

The other pressing concern raised by medical professionals is that, due to uncertainty, the predicted outcome can simply be mistaken. Such an erroneous test result is called a *false positive* in medicine. If life-sustaining treatment is withdrawn based on a false positive prognosis (or a lack of prognosis), patients who could have recovered with little neurological damage if care had been continued, now die. This is an undesirable outcome that all decision-makers I spoke with wish to avoid. However, as we will see in chapter 6 and 7, it is really hard to avoid perpetuating such mistakes because false positive cases cannot be discovered afterwards. Once the decision to withdraw life-sustaining treatment is made based on the prediction that the outcome will be poor, the outcome becomes poor. Reinhart's intensivist commented aptly: The difficulty is that, as intensivists, we're almost always proven right. 'Cause if we stop, they die, right?' While these self-fulfilling prophecies make the risks regarding withdrawal of life-sustaining care hard to detect, it is widely acknowledged that the risk is real (Wilkinson 2009, Geocadin et al. 2012). Recently, several anecdotal but significant reports on false positive findings in prognosis of postanoxic coma have been published (Maciel et al. 2017, Amorim et al. 2018).²

While self-fulfilling prophecies are caused by uncertainty regarding what outcome can be achieved, there is clearly also uncertainty about which outcomes are acceptable and which ones are not. Both uncertainties are problematic and both play a role in each patient case. Ultimately, decision-makers must act knowing that withdrawing life-sustaining treatment leads to certain death, whether that death was inevitable to begin with, whether it prevented a worse outcome, or whether the death should have been avoided. As a result, both the lack of prediction as well as its inevitably probabilistic character (Smith et al. 2013), including the risk of false positives, creates a desire for more prognostic certainty. With such risk to consider, one's first response might be to not withdraw in cases of uncertainty. Continuing life-sustaining treatment, however, holds another great risk.

4.2. Choosing to continue treatment

Whereas WLST typically leads to one outcome only, continuing care can have a variety of results. In what follows, I juxtapose two scenarios in which treatment was continued regardless of prognostic uncertainty, yet with widely different outcomes. For both Vanderbilt and Felts, established neuroprognostic tests produced inconclusive results while cEEG indicated poor outcome, and both patients survived. Although care was continued, both families were told to expect the worst and both nevertheless kept hoping for the best.

² At the same time, recent postmortem investigations, albeit comprising limited samples, show no false positives. (van Putten, 2019, Endisch et al, 2020)

Vanderbilt's family was rather optimistic and set on maximizing care. His wife, Johanna, explains their motivation:

The tests were really pessimistic, but because there continued to be triggered responses, the next step was always just keep going, keep going. Maybe also because of our behavior. We really wanted to go on [...] Because you still think, you also hear, also in the conversations it was said 'it is possible that someone wakes up from a coma like this, saying 'hello, when can I go home?' But it can also be that someone is vegetative or something like that. They told us that too. And you also know this rationally, but then I still think, the chance that he is the one that likely says 'Hello, can I go home?' Yes, the chance exists, so keep going. And perhaps, it's because I could still hear him on the phone (the last thing I heard him say) 'Where are you? I need you.' (starts crying) If he says 'I need you' you do everything (pauses) and you continue to do that, without hesitation. (keeps crying)

When Vanderbilt finally regained consciousness, his eyes were open, he could be put in a chair but was rather unresponsive. Johanna testifies:

I still have this image in my head when I saw him for the first time, next to the bed, in a chair. It was him sitting there, sure, but... then I thought, I had a flash of 'oh, how awful. So this is how it could stay... He could remain in this state.' That, he would not have wanted. [...] When I saw him like that, I thought, 'this cannot be the end'. But I also told myself, 'he's in a recovery phase, so it'll get better'.

Vanderbilt's daughter, Roos, offers a similar perspective regarding the first months after her father regained consciousness:

At one of the nursing homes, in the very beginning, I cried so much. I thought, how awful! This man can't do anything anymore, he just sits there. He can only lie down and wait till someone comes to him, to go eat, or drink, or entertain him. It was so overwhelming.

As we shall see, Vanderbilt's family, six years after his cardiac arrest, still faces some of the gravest consequences that can result from continuing care in cases of high uncertainty. The hope that pushed the Vanderbilts to go on, however, is the same hope that the Felts' family held onto.

Anke, Felts' wife shared her recollection of the 'bad news' conversation with the intensivist and neurologist:

When the doctors left that little room, we were all in tears and upset and grieving. And I don't know how it happened, I took her hands or she took

mine or we did simultaneously, but the nurse and I, we looked each other straight in the eyes and she said, I don't remember exactly, but 'you can always keep hoping, keep trusting' something like that. That moved me so much (voice trembles). Really. The fact she said that after the tremendously bad news... We were absorbed by how small the chances were and she knew that too but I thought, I will hold on to that. And we did, my daughter and I both.

While remaining hopeful, the Feltses expected the outcome to be grim. Yet Felts recovered well despite the prediction and has since gone back to working full time. He recounts what he remembers from his awakening:

[When I woke from the coma] my brother was standing in front of my bed. He had come especially from Indonesia because he had heard, obviously. They had told my wife and daughter that it was 90% certain that it wasn't going to be okay. When the doctors left, there was a nurse who said, let's first wait and see how he comes out. And, yeah, I woke up and the first one I saw was my brother—his name is Jake. I think I was convinced that I was laying in my own bed at home and that my brother was standing in the bedroom (laughs) so I said 'Jake!? What are you doing here?' So he starts crying. He starts calling my wife, saying 'he recognizes me'. That was a very beautiful moment. After that, I don't remember much. But that's when I woke up.

Anke was on the phone, the landline, when Jake called her cellphone:

I was on the phone with an old friend but I ran around screaming. I was really screaming. Oh, he recognized Jake! We immediately went to see him, but he didn't recognize me. That felt so strange. How is it possible he recognizes his brother, but not me? I was worried about the cold I had and so I was wearing a cap in front of my mouth and face. And I was wearing glasses he later said he didn't recognize. I had had them for a while but to him, they were new. A few hours later, the nurse told me to take the cap off; that it was much more important that he recognizes me. When I took it off, he recognized me! That was such a special moment. We held each other and I told him I loved him and he said, 'I love you too'.

Although the outcome of Felts and Vanderbilt couldn't have been further apart, the uncertainty their families dealt with was the same and the experimental prognoses were equally pessimistic. Both shared the same hope, the same fears. Jolien describes a moment before her father recognized his brother:

One of the nurses said to me, 'Oh, this must surely, really, be very hard. To see your father like that?' And I thought, 'Oh, gosh, if you say this now already...' [...] He wasn't doing much at that point. You could go see him,

but that was it. He looked at me, but he didn't see me, I think. His head moved but his eyes didn't. So he looked in my direction with his head but--with his eyes--he looked straight through me. And when she said that, I thought, 'right, it might just stay that way'.

In retrospect, continuing care for Felts was the right choice. While the uncertainty regarding the eventual outcome turned out to be significant (seeing his actual outcome was very different from what the prognostic information had indicated), there is no uncertainty whatsoever that the actual outcome was desirable. For Vanderbilt, it's complicated. Vanderbilt's son, Natan, says that if he had the chance to go back in time—knowing how things turned out in the end—he would want to stop treatment in the ICU. Yet, being back without that knowledge, he would probably still share his mother's and sister's desire to go on. The differences in valuation of the actual outcome prove that the Vanderbilts disagree about whether or not the value of the eventual outcome justified continuation of treatment. Yet, without certainty that the actual outcome would be unacceptable, even Natan would rather face the uncertainty that continued treatment entails.

These two stories teach us how decision-makers experienced the option to continue treatment in the face of uncertainty. While both families shared the hope of functional survival that drove the continuation of care, the alternative would have led to certain death. In both cases, there was not just uncertainty about what outcome could be achieved. In addition, there was uncertainty about what outcomes would be acceptable and for whom.

4.3. Delay or advance decision-making: extended and reduced windows of opportunity

The experiences discussed in the previous two sections show how both withdrawing and continuing life-sustaining treatment come with their own risks. They vividly illustrate how decision-makers in clinical practice face a dilemma in which there is no easy way out. However, the intensity of the dilemma depends, to some extent, on the timeframe in which WLST is likely to result in the patient's death. While this wasn't immediately apparent in the Dutch ICU's I observed, interviews with medical professionals in acute and long-term care and empirical observations in the Netherlands and abroad showed that, depending on the length of said timeframe, decision-making can be delayed or advanced.

Let me explain. The aim of WLST is to allow a patient to die a natural death. There are biological limitations, however, to the period in which this aim can be realized. Coma is generally expected to last two to three weeks although the exact time can vary from patient to patient (Jennett & Bond 1975, Katz 2009). After this period, patients may be able to maintain bodily functions such as breathing even after life-sustaining support like a ventilator for mechanical breathing is withdrawn.

This means that withdrawal after two to three weeks may lead to survival with severe disorders of consciousness, like vegetative state or minimally conscious state. Therefore, the temporal window during which withdrawing a patient's life support can be expected to lead to the patient's death is commonly called the *window of opportunity* (Wilkinson 2011, Kitinger & Kitinger 2013).

However, the length of this timeframe is not constrained solely by biological factors but by political ones too. The comparative empirical investigations I conducted in the United States and Israel made this especially clear. They showed that the end of the window of opportunity does not necessarily coincide with the end of coma. Windows of opportunities can be shorter or longer depending on local legal policies. In Israel, for instance, life-sustaining care after intubation can no longer be withdrawn if it were to lead to the death of the patient (Steinberg & Sprung 2007). Here the so-called window of opportunity is very brief: it closes the moment the patient is intubated and admitted to intensive care. As a result, prognostic tools like cEEG used at a later time lose most of their value, whereas prognostic technologies to help decide whether or not to stop reanimation during emergency care are investigated (Einav et al. 2011). In contrast, in some US states and in countries like the Netherlands, legal measures exist to end a person's life even after a person can breathe without the help of the ventilator, thus extending the window of opportunity indefinitely.

The dilemma and the psychological pressure it entails can thus be alleviated by longer time frames or intensified by shorter ones. One could argue that, in countries where treatment can be withdrawn for patients with disorders of consciousness, there is no real window of opportunity since there is no point where it is 'too late'. Indeed, the Dutch expert on disorders of consciousness I spoke to confirmed that there are sufficient legal measures for a patient in the Netherlands to not have to go on living in such conditions. My observations in the ICUs in the US and the Netherlands showed, however, that decisions in the Netherlands are made relatively early. The US practitioners I spoke to said they usually wait until one week after admission to the ICU, but it is not uncommon to wait much longer than that. Whereas the Dutch intensivists I spoke to said decisions at 72 hours are common in the Netherlands. This practice is furthermore in line with the Dutch prognostic guidelines (Horn et al. 2019). Only when no prognostic indication is available, do they wait longer but commonly no longer than one week. Thus, my empirical investigation shows that, even though decision-making could be postponed in the Netherlands, it often is not.

Based on the conversations with medical professionals, there are several considerations why 'waiting it out' is not more common in the Netherlands, regardless of the extensive window of opportunity. First, as discussed previously,

if there is an indication the patient won't recover well, like with Reinhart, continuation of life-sustaining treatment is considered as a prolongation of the death process and the patient's needless suffering. Second, even if patients with disorders of consciousness can be allowed a natural death, that still leaves a small minority of patients for which the window of opportunity continues to exist and for whom the consequences can be disastrous. Patients like Mr. Vanderbilt regain consciousness in a way that no longer permits withdrawal (if that were desirable). This means the window of opportunity closed and will have closed even in countries with the largest windows of opportunity, like the Netherlands. For such patient cases, the dilemma remains. Furthermore, decisions to let go of a patient at a later stage are emotionally and psychologically challenging for both the patient and their family. Finally, continuation of life-sustaining treatment in intensive care is very costly. The Netherlands has a universal healthcare policy which requires decisions in the distribution of limited resources. National withdrawal policies that keep the window of opportunity rather short for patients with little to no chance of recovery (Royal Dutch Medical Association 1997, 2005; Gerritsen et al. 2009; Dutch Ministry of Health 2020) make sense when aiming to distribute limited resources. I will return to the relation between windows of opportunity and resource distribution in the next section (§5.3.).

In short, decision-makers in the Netherlands experience the clinical dilemma due to prognostic uncertainty very strongly, even though the dilemma could in theory be alleviated in other ways. This section shows how time pressure makes already difficult decision-making more difficult and how there is uncertainty even about how quickly decisions should be made.

5. BRAVING UNCERTAINTY: AVOIDING THE GREATEST HARM

We have seen the different ways in which decision-makers can respond to the dilemma. Investigating the respondents' reasoning shows that when physicians and next of kin deliberate about which way to go, continuing or withdrawing life support, they generally seek to avoid the greatest harm.

Three types of potential harm weigh in on their decision: First, the patient's potential suffering is weighed against the patient's perhaps avoidable death. Second, the emotional and psychological impact of each scenario on the family is taken into account. And third, although not usually made explicit in deliberation of individual cases, the societal cost in light of distributive justice is deeply rooted in national policy and medical practice. In what follows, I use the outcomes of the different cases to illustrate why these considerations are essential when attempting to avoid the greatest harm but nevertheless challenging due to the risk caused by uncertainty.

5.1. The suffering of the patient versus their life lost

Reinhart died after life-sustaining treatment was withdrawn. No one can be absolutely sure whether or not he could have survived or to what extent he could have recovered. Based on the prognosis, however, it is likely he was spared a terrible fate and his family says that, in light of the prediction, he would not have wanted them to risk continuing life-sustaining treatment.

Reinhart's wife, Irene, explains part of her reasoning:

For him, what he absolutely didn't want—I think he would be very angry with me if I had kept him alive but that he wouldn't be able to independently go to the bathroom, wouldn't be able to talk, making eye contact, if he wouldn't be able to see the kids growing up, just hear stuff. No, I think that, although this is very hard, this has been the right decision.

Based on their testimonies, Reinhart's family was clear on what an unacceptable outcome entailed for him and based on that, they were not willing to risk continuation of care if it meant risking an outcome like, for instance, Vanderbilt's.

Mr. Vanderbilt initially survived with relatively good motoric function. He could, for instance, learn to walk and get dressed. His cognitive abilities, however, remained extremely limited. He can hardly speak, he has difficulty understanding what is said to him and misinterprets or momentarily forgets common social practices, like whether or not going to the bathroom is a private matter and where it takes place.

Based on observations of Vanderbilt's interaction with his family and their testimonies, one of the factors that seem to make Vanderbilt's situation so heartbreaking for himself and his family is that he is often sufficiently conscious to be aware of his condition. In the past, the sudden realization of his situation and the confusion and stress that came with it, triggered panic and anxiety attacks.

At times, these could lead to obsessive behavior like asking over and over where the bathroom is even when he is already brought to the bathroom, looking at the toilet, but no longer understanding what he is looking at or what his interest was in finding it. Considering the family knew Mr. Vanderbilt as a devoted husband and father, a responsible and highly functioning project manager, and the head of his sports team, they are inclined to interpret his desire to know where to find the bathroom as a conscious effort to avoid an embarrassing incident in the future. His awareness of his limitations and the lack of short-term memory can thus result in a desperate attempt to hold onto a thought that he feels deep down is important.

Other times, the apprehension of his situation followed by panic and anxiety attacks, led to hours, sometimes days and nights of weeping and severe depression which, ultimately, had to be remedied with medication. The expert on disorders of consciousness who was asked about his condition called it *existential suffering*. On a few occasions, the anxiety and panic attacks evolved into uncontrollable rage and violence and, one time, both family and healthcare practitioners had no measure left to them other than to hold Mr. Vanderbilt down and completely sedate him.

Roos explains,

When my father had to be tied up a year later, that hit me really hard. It was so inhumane. Like the Exorcist, it was like that. Tied up and making crazy faces, sticking his tongue out, like a monster. The whole hospital could hear him scream. They used so much medication; he was in a coma for days. But they had to, he was so strong. Even with all the medication, five people couldn't hold him down.

His wife, Johanna, when she recalls the same event says, 'It cannot escalate like that again. I can't have it. It can never happen again, what happened then...' then she starts crying.

Crucially, when reflecting on what harm to avoid, many respondents immediately assume that remaining unconscious (in vegetative or minimally conscious state) is the worst outcome, to be avoided at all cost. Yet, all of my respondents also agreed that having sufficient cognitive capacities to be aware of your limitations would make the experience much more excruciating, especially when the awareness seems to fade in and out.

Vanderbilt's children also noted that this fluctuation in awareness adds difficulty to knowing when to take his personal assessment of his situation seriously and when not to. And, as Roos testified, the stakes can be of vital importance.

There was a moment where he would repeat on and on 'want to go home, want to go home, want to go home.' And I tried to explain to him that it wasn't possible, or something like that, I don't remember exactly what I said, but he answered, 'then I want dead'.

However, even if there is clarity about what quality of life a patient considers unacceptable, certainty about which outcome to avoid may not be enough to be able to decide whether or not to withdraw treatment. As I have pointed out above, in the process of attempting to avoid an unacceptable outcome, there is a small chance that the patient may actually lose a very meaningful outcome in the process. As long as any uncertainty remains about what outcome can be attained, the dilemma remains.

If Reinhart's prognosis were incorrect and he could have survived like Felts did, would that have been worth the risk?

The problem with comparing the different harms to the patient in each scenario is that in one case, something very tangible is happening to the patient whereas in the other case, this is missing. Think of how both the value and long-lasting effects of Felts' survival and the detriment of Vanderbilt's suffering can clearly be observed. However, when a patient dies after WLST, the good or harm done by preventing either someone's suffering or their recovery remains unknown. There's no way of knowing whether an essential value or devastating harm was taken away. One can ask Felts to imagine what it would have meant to him to miss out on the years of life he gained upon his recovery. But how does one compare the imagined harm to a patient due to the loss of their life, once they no longer exist to experience that loss, with the very experienced harm of surviving with severe neurological damage?

Yet, there may be an even more significant asymmetry between the ways people evaluate harms and benefits. Even when Felts praises the value of living beyond measure, would the possibility of such invaluable benefit outweigh the threat of immeasurable suffering? In *Harm and its Moral Significance*, Shiffrin states that 'morally, the aversion, prevention, and relief of harms seem, at least presumptively, to matter more than the provision, protection, and maintenance of comparable and often greater benefits' (2012, 358). Whether or not this is right, this is precisely the asymmetry clearly visible in Dutch practices where a choice has to be made between risking a terrible outcome or risking the loss of a valuable one. When faced with the risk of great harm or loss of great value, precaution creates a bias that favors the avoidance of harm.³

As such, my empirical investigation confirms that, as a result of this asymmetry and the discrepancy between harms and benefits that are visible and experienced and those that are not, the latter easily fade into the background. This imbalance is exacerbated over time by the fact that, due to self-fulfilling prophecies, one cannot find out in retrospect whether such harm actually took place. In chapter 6 and 7, I discuss this issue in greater detail. Thus, in cultures where sanctity of life is not necessarily trumping quality of life, it makes 'sense' that, in practice, the suffering of the patient influences decision-making more strongly.

5.2. Emotional and psychological impact for next of kin and caregivers

Decisions about WLST in the face of uncertainty do not just have far reaching consequences for the patient. The family, too, is greatly impacted, for better or worse. Either a family loses their loved one and, while this is bad enough, they may never know whether that loss was inevitable, perhaps even desirable, or avoidable.

³ Similar tendencies can be observed in policies responding to the 2020-2021 global pandemic.

Or the family's loved one survives but that outcome may be joyous or regrettable. Family members' reflections on the outcomes show why considerations regarding impact on next of kin and caregivers are both important and challenging.

Mr. Vanderbilt and his family underwent several additional crises after his initial recovery which led to the increasing deterioration of both his cognitive as well as his motoric skills. While his wife and children remain extremely dedicated to provide him with the best care, the emotional and psychological cost cannot be underestimated.

Describing the pain and struggle the Vanderbilts face on a daily basis would take us too far afield and it is questionable whether any account could actually do justice to their experience. My aim here is only to give a glimpse of the extent to which the situation has impacted their lives.

Roos said about the impact on her partner and their relationship:

Last year, my relationship also ended. We've been through so much together. He was constantly there with me, from the beginning. The whole situation has been really, really hard for [my ex] too. At some point it costs so much energy. I could tell, I saw how it affected him, it was ruining him. He was so very much involved and he crashed hard because of it. I hope he recovered now, but just like my mom, my brother, and me – we've all really suffered from this.

Both Vanderbilt's wife and daughter have severe health problems due to years of continuous, high level stress. The daughter has been struggling to manage a very serious and visible stress-related skin problem. Her brother Natan worries that their mother, who he says is completely worn out, is going to collapse and they will end up having to nurse two people.

Johanna admits that, at some point, it becomes a matter of whether her body can handle it, because she has significant physical problems that she was told are due to emotional and psychological stress.

What I've been living now for three and a half years, I can feel that it's become a... a burden is such a strange word, an unpleasant word, but it has so much impact on your daily life.[...] I still cry a lot, even though I feel like I didn't suppress anything. It just rules my life.

While she and Roos see some value in Vanderbilt's survival, Natan concludes that the situation has become detrimental to everyone involved.

I don't want to end on a negative note but I hope my father will soon find peace, that he'll pass away in the coming time. How strange or how harsh that may sound, I do hope that. [...] I have a wife, two small children [...] For the synergy of all, it would be better if my father were no longer there and everyone could start picking up the pieces and rebuild themselves.

He emphasizes that, although the best way forward, his father's eventual passing will not be easy:

I'm afraid for what that situation will bring because you've got to go through that thick, bitter pill, one more time. You've got to take one more serious hit, but then you can rebuild. For my sister and mother, it would be better too, absolutely. I'm 100% convinced of that.

The toll Mr. Vanderbilt's survival has on the rest of the family is clearly great and it is that kind of pain and distress that the physicians I spoke with often pointed to when they evaluated potential outcomes. The experience clinicians have with such grave outcomes compels them to protect the patient but also the family. The higher they estimate the risk of an outcome like Vanderbilt's, the more they are inclined to advocate for withdrawal.

Yet, again, the very visible and excruciating costs of erring on the side of survival stand in stark contrast with the invisible cost of the death of a patient who, potentially, might have survived with an acceptable quality of life. Although Felts' wife and daughter remain fearful of losing him, they are overjoyed to have him back after having faced the actual prospect of his death. Had life-sustaining treatment been withdrawn based on the prediction, the error would have never been known. In such cases, much depends on how the family experiences the decision making.

All patients' families concurred that when they felt that their loved one was well cared for, that they were listened to, and that the healthcare professional did their best to make decisions with the wellbeing of the patient in mind, they were more likely to have little remaining discomfort about the uncertainty, even when they were aware of possible errors. They understood the risk as being inherent to the practice and trusted practitioners were doing the best they could. Once they accepted the decision as the best possible way to do justice to the patient, they seemed unlikely to question it later. Especially when they were given enough time to process.

Reinhart's wife explains how much time helped her:

You had a week to say goodbye. It's still extremely hard but I did think, I still have him, I can hold him, still go through his hair, I can still scold him,

like 'damn you, what did you do, why didn't you go to the doctor sooner, why didn't you listen, or, this was your way,...' I am so happy I received six, seven days. Despite that within 48 hours I think they pretty much knew already that he wouldn't ever be himself again.

Irene felt she was given additional time beyond what she was owed following protocol. Even if this was not necessarily the case, such interpretations and gestures of care and compassion are invaluable. They strengthen the relationship between family and the medical staff and facilitate acceptance of the decision.

Most physicians I spoke with are well aware that establishing rapport, trust, and good communication are key to reaching agreements with relatives on what is good care and making sure relatives feel good about decisions made, even under uncertainty. Intensivist dr. Harte even makes it a priority to have the first interaction with family, whenever possible, for precisely that reason.

The first interaction is so important, you cannot miss it. It's your chance to get the family's approval. If you make sure that that first contact is good, then you win the family's trust and you start out with ten points. Even if then someone messes up a little along the way, you have some wiggle room because you started with ten points.

What became increasingly evident throughout my fieldwork is that, when there is no certainty to be had, it may be more important to make a decision that everyone affected can live with. This is an important shift. Once the call for more certainty about the outcome is left unanswered, the most urgent need is to better understand 'what exactly it is everyone can live with'.

The acceptance of risk notwithstanding, the cost of losing a loved one altogether, whether preventable or not, should of course not be underestimated. Reinhart's death has devastated his wife's and sons' lives in the past year even though they felt good about the decision. Similarly, had Vanderbilt's care been withdrawn, the family would still have suffered a great loss. The pain of loss and the mourning process that follows is likely to be devastating to any family. However, it is a very common and familiar devastation that society is better equipped to deal with, simply because death is an inevitable part of life, for all.

The loss experienced by families like the Vanderbilts, is much harder to understand for outsiders and more complicated, in part, because it is less understood. Institutional support is lacking and the coping process can be much longer. When the patient finally dies, the family will still have to go through a mourning process. 'That thick, bitter pill', as Natan said. In the meantime, however, they have suffered

a loss that is harder to mourn, since their object of affection is still very much present. Yet, as Jolien summarized: 'The one you had is no longer there'.⁴

A telling indicator of such psychological challenges was when Natan misspoke during the interview. He explained that he couldn't stand looking at old pictures from the time his dad was still healthy. 'I'm in pain when I think of the situation before the cardiac arrest (pauses) because then I know what it was like and what I could've had if he were still alive. (coughs) So the missing hurts me more than what is happening now.' Natan did not even realize he had spoken of his father as no longer alive until it was pointed out to him. After I told him, Natan responded that that showed exactly how he experienced the situation.

His mother, in contrast, although she admits to feeling she has lost something, says she gets angry when people talk to her about mourning. She does not want to hear the word *mourning process*. 'Yeah, I don't want to hear that word. How so, mourning process, my husband is still here, so.' Although different relatives experience the challenges differently, all family members experience a great impact on their mental and physical health.

In sum, whereas continuation of life-sustaining treatment can lead to great joy (as it did for Felts and his family), it can cause a lot of pain and detriment too (as it did for Vanderbilt and his family). Withdrawal of life-sustaining treatment, on the other hand, always leads to death and a more common mourning process. Crucial to that mourning process then is the way in which the decision to withdraw is made, whether the wishes of the patient and family were taken into consideration, whether the family felt heard, and whether they felt the healthcare professionals had the patient's and their best interest in mind. My fieldwork shows that the decision's impact on the family is important and clinicians feel it should be taken into account in the decision-making. The families' testimonies show *why* this is important. It nevertheless remains challenging to understand how that interest is best considered in each individual case, especially considering the unknowns.

5.3. Societal cost and distributive justice

The previous two sections describe why considering the impact on both patients and their families is relevant but nevertheless challenging. Yet, considerations beyond the individual, like societal costs inevitably weigh in as well. Although my observations and interviews with medical professionals confirmed that economic cost is usually not explicitly factored in when considering WLST for a specific patient, my findings also show that the clinical practices are inherently restrained

⁴ Studies on family and caregiver experience with grief regarding dementia, Alzheimer's, and disorders of consciousness are increasingly providing insight in these mourning processes (Chan et al. 2013, Blandin & Pepin 2017)

by resource limitations. In what follows, medical decision-makers discuss to what extent resource concerns do and do not affect them and when they are considered more or less important in the decision-making process.

While non-medical respondents do recognize the need for limitations in care, most feel uncomfortable expressing judgments about what limits to impose on others. Furthermore, in their personal deliberations, considering cost seems almost unthinkable. This position is also well-reflected in our survey among Dutch citizens, presented in chapter 5. In contrast, physicians seem to consider how to distribute limited resources more frequently and in increasingly less abstract ways. Those I spoke with feel that hospital management informs them more and more about specific costs.

Dr. Valk:

You're increasingly considering cost. It was always somewhat in the back of your head, but now it was made more explicit what [for instance] lab work costs, and that really makes you consider not doing it as often, 'cause it *does* save money. If you have two and half thousand admissions a year, that can make a huge difference.

These specific considerations differ from what Reinhart's intensivist refers to as 'always in the back of your head'. Distribution strategies in the form of 'triage' where patients are sorted into 'immediate', 'urgent', 'non-urgent'—although originally designed for situations of mass-casualties—are now common practice (Robertson-Steel 2006), especially for intensive care units. So rather than allowing a patient to die simply because they have poor chances, the rationale is that a patient with much higher chances gets the priority. Here too, one can recognize a motivation to minimize harm on a collective level. In countries like Israel, where national policy requires that patients are kept on life-support even when their chances of regaining consciousness are zero, the strain on resources is strongly felt and triage practices at the emergency department and ICU have become inevitable. In contrast, these kinds of considerations in case of individual patients are rarely required in Dutch practices.⁵ There, withdrawal policies for care that does not fall under 'zinnige zorg', literally translated as 'sensible care', allow for more availability of resources.

Clearly, having resource distribution dealt with in national policy relieves some of the pressure on the ground. In fact, when asked about the effect of cost on postanoxic care, dr. Valk adds that his attitude changed in the other direction.

⁵ In extreme circumstances like those experienced during the COVID pandemic, this kind of pressure on ICU hospital beds is felt even in countries like the Netherlands, precisely because they are not equipped to deal with that many patients needing life support.

Whereas before he was keeping to the guidelines more strictly, withdrawing sooner, he has become more careful over the years. 'Due to the uncertainty', he says, 'I know that I might be stripping someone of a viable chance'.

I'm not in so much of a hurry anymore. Now I think, okay, so it'll take a little longer. I don't mind. But those were the guidelines, within 3 days, otherwise someone spends another week in our ICU while they won't wake up anyway, takes up space, costs money, there's uncertainty for the family,... so following protocol, you're supposed to know after 3 days.

While Reinhart's intensivist chooses to wait longer in cases of uncertainty, the three-day guideline is on the very short end of the spectrum. In the US practice I observed, even though it is also set in a teaching hospital with physicians conducting research on cEEG for early prediction, it is common to wait a week, sometimes weeks in cases of uncertainty. And even there, the neurologist heading the research on cEEG said 'I feel we withdraw too soon'.

As previously mentioned, the ability to delay the decision to withdraw depends on the window of opportunity. Comparing practices in countries with widely different policies brings to light the direct correlation between the window of opportunity and resource availability and how decisions on WLST are shaped by their larger contexts. Typically, longer windows of opportunity would delay decision-making and thus increase treatment times, leading to higher costs while shorter windows would lead to faster decisions to WLST, and thus fewer costs. However, practices that have no real *window* as such—either because WLST is prohibited and thus never possible or because WLST is always possible—run up against limits. Windows that are too long or too short seem to demand practical or policy interventions in the opposite direction.

For instance, the Netherlands, as explained previously, could have a window that never closes. Put into practice, however, such a policy would lead to extremely high costs of care on the one hand and, on the other hand, many patients with severe disabilities who subsequently need to be let go after they have ended up in nursing homes. Such consequences can be remedied by artificially shortening the window of opportunity and allowing WLST following a prognosis of poor outcome or an estimation that further treatment will not be in the best interest of the patient.

In contrast, the empirical study I conducted in Israel shows that in places where WLST is forbidden once a patient is intubated, the window has closed before it could open. With all patients remaining on life support, the pressure on resources is intensified. Here too, in order to alleviate both the costs as well as the suffering of patients surviving in poor neurological state, a very short window emerges in practice early on, arguably too early, at the level of the emergency services. This

shows that making decisions about WLST as such is shaped by the context in which the practice takes place⁶.

This may explain why all the intensivists and neurologists I spoke with, regardless of the country, said that, in general, they do not explicitly consider societal cost in their decision-making about a specific patient. And yet, broader structural conditions clearly shape the way the dilemma is experienced. Cost, therefore, seems to automatically weigh in on the decision through national policies, institutionalization and much depends on the healthcare and insurance policies. In the Dutch national guidelines, especially, equal distribution is very much institutionalized.

Dr. Harte elaborates:

If you get oncological medication against lung cancer, it can cost 140000 euros a year. While the Dutch Care Institute (Zorginstituut Nederland) has defined that one human life year may cost up to 80000 a year. Therefore, says the Care Institute, this medication cannot be included in standard care.

Yet, when asked about patients who would be able to pay for it out of pocket, the intensivist responds:

Frankly, I'm actually not entirely sure I would want to give that treatment. (pauses) Because if my neighbor who is but a carpenter has lung cancer, I would then want to give him treatment as well.

His response shows that ideas of equal distribution are internalized too. The intensivist admits, though, that this internalized sense of equal care is limited. Remaining consistent with regard to equal distribution across national borders would drastically change the way medicine is practiced. Even so, this is not something a single practitioner feels they can influence.

Vanderbilt's intensivist continues:

If I had something with which I could give everyone equal care in the entire world, then I would of course immediately do that. It would create a terrible dilemma though. I would, for instance, be the first to close this entire [intensive care] unit. Because it makes a lot more sense to distribute [basic medicine] and build in new knee prostheses than having people admitted to intensive care.

⁶ In-depth historical investigation is required to better understand and fully capture the developments in countries with either very permissive guidelines (like the Netherlands) or very restrictive guidelines (like Israel).

On an international level, equal care would indeed require extensive redistribution. This would mean that many treatments, like intensive care, would become 'too expensive' and the standard of care could be significantly reduced for people in affluent countries—eradicating the dilemma in prognostication of postanoxic coma altogether. But even in affluent countries, when confronted with finite resources, as seen during the global pandemic, avoiding the greatest harm may call for practices of triage where patients most likely to survive will be chosen to be treated over more desperate cases. Ultimately, the tendency to avoid the greatest harm on a societal level is intrinsically present in the decision-making process.

6. CONCLUSION

Felts' neurologist, while giving voice to the weight of the dilemma decision-makers face, summarizes the burden of having to make decisions about life and death in the face of uncertainty.

With patients with a severe encephalopathy, both deciding to continue and deciding not to continue is extremely loaded. [...] You can imagine as a researcher or a bystander that this is very hard but the moment you really carry the final responsibility... yes, it elicits feelings of uncertainty, of course it does. Is it *good* to do this now?

In answering that question, medical professionals but also patients and relatives, seem concerned not just about *what* outcome can be expected but also *how* this outcome will be valued.

I set out to investigate how decision-makers in the context of postanoxic coma experience the uncertainty regarding neuroprognostication and, subsequently, how—in light of that uncertainty—different concerns weigh in on their decision-making process. First, I described the medical dilemma caused by prognostic uncertainty and the choices decision-makers are faced with; withdrawing or continuing life-sustaining treatment, or, in some contexts, advancing or postponing the decision. I illuminated what considerations are taken into account and how.

I showed that, when facing uncertainty, deliberate decision-making is curtailed by a maximum of harm that decision-makers are willing to risk and limited by cost considerations that are inherently present in clinical practice. Therefore, the concerns that emerge in practice seem understandable: each decision requires careful consideration of potential harms to the patient, the patient's family, and society as a whole, as these have far-reaching consequences long after the decision is made. All of these considerations, however, are strongly dependent not just on *the reliability of the predicted outcome* but on whether that *predicted outcome*

correctly projects the kind of harms and benefits it will bring the patient, their family, and society and whether these will be regarded as acceptable or not and by whom.

In essence, analysis of the decision-makers' concerns and how they weigh them in practice show that the avoidance of potential harm for a particular patient necessitates answers to two very different questions: *what outcome can be achieved* and *will that outcome be acceptable*? While decision-makers struggle to find answers to *either* of these questions, efforts to reduce neuroprognostic uncertainty are focused on answering only the first question. Innovative technologies are developed to help reduce uncertainty about what outcome can be achieved but the challenges with regard to whether a projected outcome is acceptable have been neglected. The quest to reduce neuroprognostic uncertainty will require further investigation regarding what outcomes are found acceptable, how, and by whom. Special attention is paid to this last point in the next chapter and in chapter 5. The following chapter offers a deeper analysis of the two questions that emerge in practice and the answers they require.

Chapter 3

Predicting Medical Futility After Cardiac Arrest: pulling apart physiological and normative uncertainty

ABSTRACT

Deciding whether to withdraw life-sustaining treatment from a patient in coma is typically motivated by medical futility judgments. In the medical futility literature, distinctions are made between physiological and normative (or quantitative versus qualitative) futility. Uncertainty regarding physiological futility is highly acknowledged and neuroprognostic tools are developed to address it. Yet, my fieldwork shows that normative futility judgments are often just as uncertain—not just due to uncertainty about the physiological outcome but also due to uncertainty about how the outcome will be experienced and evaluated. Additionally, there is uncertainty about how environmental, contextual, and institutional support can impact said experience and evaluation. I argue that any serious quest to reduce neuroprognostic uncertainty requires distinguishing between *uncertainty whether the proposed treatment can physiologically achieve the desired effect (physiological uncertainty)* and *uncertainty whether the physiological effects of the treatment constitute an acceptable outcome (normative uncertainty)*. I conclude that normative uncertainty can render normative futility judgments so undetermined that it is more accurate to say they require prediction rather than identification. Hence, neuroprognostic practice may require two—not one—predictions and, as a consequence, different parties may offer distinct expertise to inform such distinct predictions.

Predicting Medical Futility After Cardiac Arrest: pulling apart physiological and normative uncertainty

1. INTRODUCTION

In increasingly more countries around the world, withdrawal of life-sustaining treatment (WLST) is legally allowed when a case can be made that continuing such treatment is medically futile. Due to prognostic uncertainty, however, assessing the futility of continued treatment is far from obvious. Meanwhile, continuing futile treatment may lead to unacceptable outcomes. As a result, a critical question arises in medical practice: whether or not to withdraw life-sustaining treatment from this particular, individual patient? In particular, professionals on Intensive Care Units (ICUs) have to face this question regularly, for example when patients are in coma after cardiac arrest.

As previously mentioned, researchers and developers of new prognostic tools expect such tools can reduce the prognostic uncertainty, and thus help to decide whether or not continued treatment is futile. Before we can evaluate whether better prognostic tools can help reduce uncertainty, however, we need a good understanding of what uncertainty is at stake and what causes it. In the previous chapter, we saw that both medical professionals as well as the patient's family struggle to answer two very different questions: *'what outcome can be achieved'* and *'will that outcome be acceptable'*? These questions are similarly reflected in the bioethics literature on medical futility where they reoccur in terms of physiological versus normative or quantitative versus qualitative futility. Aiming to better predict what neurological outcome can be achieved, neuroprognostic innovation is mostly focused on physiological futility. In this chapter, I will take a closer look at the uncertainty decision-makers face regarding normative futility when they consider whether a given outcome will be acceptable.

For background, I first introduce the notion of medical futility and formulate clear and comprehensive definitions of medical futility and utility judgments, distinguishing between normative and physiological judgments (§2). Second, I discuss how, in neuroprognostication, the two types of judgment are compounded in varying ways and degrees, all of which imply that normative judgments are more or less determined and need only be identified correctly (§3). Third, I show that some of these normative judgments are especially unstable and, therefore, undetermined (§4). Finally, in section 5, I argue that neuroprognostic uncertainty is an entanglement of two distinct uncertainties that require two, not one, prediction (§5.1). This insight furthermore raises the question of who is in the best position

to make which prediction? (§5.2). I conclude with a summary of the findings and guidance on how to responsibly move neuroprognostication forward in both research and care (§6).

2. JUDGMENTS OF MEDICAL FUTILITY

The notion of medical futility originates from medical practitioners but has been picked up in the bioethics literature. Both medical professionals and ethicists have long discussed when treatment is or is not futile and who should determine this (Youngner 1988, Lantos et al. 1989, Schneiderman et al. 1990, Callahan 1991, Pellegrino 2005, Bagheri 2008). However, these discussions have not led to a common, universally (or even widely) accepted standard for what medical futility entails or how it should be responded to in practice.

At the core of the disagreements lies intricate complexity. Judgments of medical futility typically refer to (a) medical facts, (b) values, and (c) socio-economic considerations (Bernheim et al. 2013). Although the three aspects are interdependent, Veatch (2013) argues that it is essential to clearly distinguish between physiological and normative futility, and separate both from resource allocation concerns. Førde (1998) explains, in the same vein, that medical futility has 'two important components; the *probability* and the *desirability* of reaching a particular goal by medical intervention'. While the probability of reaching a specific goal depends on physiological and technological factors, the desirability depends on normative factors.

Hence, ideally, a judgment of *physiological futility* is a *judgment of medical futility based on the observation that the proposed treatment cannot physiologically achieve the desired effect*. For example, consider a person who lost consciousness and never regains it, even when maximized care is continued. If the desired effect is to recover consciousness, further treatment was thus physiologically futile. If there is a dispute about physiological futility, it is about the effectiveness of an intervention, not about the goal of the intervention. Yet physiological futility as such, is always assessed in relation to a desired treatment goal.

Whether or not a specific physiological outcome is considered desirable, is a matter of normative futility. A judgment of *normative futility* is thus, ideally, a *judgment of medical futility based on the observation that the foreseen physiological effects of the treatment do not constitute an acceptable outcome*. Disputes about (or changes in) what is considered normatively futile are about which goal ought to be pursued, not about the medical facts regarding the capacity of the treatment. For instance, physiological outcomes after coma are typically categorized in Cerebral

Performance Categories (CPC) or following the Glasgow Outcome Scale (GOS). In neuroprognostic research, however, these categories are dichotomized in 'good' and 'poor' outcomes, indicating which physiological outcomes can be considered normatively futile. Yet, since such value judgments are hardly universal, the dichotomy has proven to be unstable. The outcome category of 'severe neurological disability' (otherwise known as CPC3) was historically considered a desirable outcome in medical practice and research until 2006. Over time, however, this normative evaluation changed from 'good outcome' to 'poor outcome' (Sandroni & Nolan 2015, Sandroni et al. 2018) meaning that treatment that results in severe disability is now considered normatively futile, whereas before, it was not.

Debates about physiological as well as normative futility are furthermore often constrained by the practical limitations of resources which limit treatment options and open the door to discussions of fairness and distributive justice. As described in the previous chapter, prognostic deliberations at the level of the individual patient, at least in the practices I observed, are primarily affected by considerations of physiological and normative evaluations of outcome, which is why I focus on those here.

A wide variety of definitions of physiological and normative, but also of 'quantitative' and 'qualitative' futility has been put forward (Schneiderman et al. 1990, Pellegrino 2005, Mohindra 2007, Sokol 2009, Schneiderman 2011). This has increased the conceptual muddle, rather than led to a shared understanding, let alone a shared *standard*, of 'futility'. Since this chapter focuses on the uncertainty that decision-makers face in practice when evaluating the futility (or utility) of further treatment, I will not engage in these discussions any further. However, I do need to clarify which terms and definitions I will use in what follows.

I am first and foremost interested in the way different parties in clinical practice make *judgments* about medical futility when facing uncertainty. These judgments furthermore imply judgments of medical *utility*¹ which are more in line with the focus on '*zinnige zorg*', literally translated as 'sensible care', present in Dutch practices. Hence, my definitions are put in terms of judgments of both utility and futility. I define judgments of physiological futility/utility and normative futility/utility, as follows:

¹ Policies allowing WLST do not always center around futility but rather it's positive counterpart, utility. In the Netherlands, for instance, the focus lies on what constitutes '*zinnige zorg*', literally translated as 'sensible care'. The difference in emphasis does tend to have consequences. For instance, shifting attention from futility to utility can change the question from 'when to withdraw life-sustaining care?' to 'when to continue life-sustaining care?' While these differences can be significant, other significant aspects remain the same. In this chapter, I focus on the overlap.

Physiological utility judgement	<i>A judgment of medical utility based on the observation (or prediction) that the proposed treatment cannot (or is unlikely to) physiologically achieve the desired effect</i>
Physiological utility judgement	<i>A judgment of medical utility based on the observation (or prediction) that the proposed treatment can (or is likely to) physiologically achieve the desired effect</i>
Normative utility judgement	<i>A judgment of medical utility based on the observation (or prediction) that the (foreseen) physiological effects of the treatment do not (or are unlikely to) constitute an acceptable outcome</i>
Normative utility judgement	<i>A judgment of medical utility based on the observation (or prediction) that the (foreseen) physiological effects of the treatment do (or are likely to) constitute an acceptable outcome</i>

Table 3.1 - Definitions of judgments of physiological utility/utility and normative utility/utility

Note that the definitions in Table 3.1 include ‘prediction’ next to ‘observation’, and probability next to fact. This is because, as we will see in subsequent sections, in neuroprognosis after cardiac arrest, all these judgments are marred by uncertainty. For comprehensiveness, however, I have included both in the definitions. This way, the definitions offered in this chapter can accommodate a broad use of medical utility and utility judgments, under varying degrees of certainty and probability.

With these definitions, I have set the stage for my exploration of uncertainty in decision-making about WLST in practice.

3. UNCERTAINTY AND NEUROPROGNOSTIC TOOLS

Uncertainty in neuroprognostication is widely acknowledged in research, literature, and clinical practice (Smith et al. 2013, Rossetti et al. 2016). In the previous chapter, we showed that, due to this uncertainty, a dilemma emerges: *Withdrawal risks that a patient with good neurological prospects dies, while continuing life-sustaining treatment risks that a patient with unacceptably poor neurological prospects survives*. Decisions to withdraw life-sustaining treatment thus remain risky as long as there is uncertainty about the outcome of the patient. Therefore, improving neuroprognostication raises hopes and expectations that uncertainty can be, if not completely eradicated, reduced to an absolute minimum thereby at least alleviating the clinical dilemma significantly. To this end, neuroprognostic tools are researched and developed to better predict the outcome of the patient. The problem is that neuroprognostic test results are then used to inform judgments of medical utility/utility² in general when even the best tools can only ever make

² For the remainder of this chapter, I will refer mostly to utility judgments alone but these continue to imply their utility counterpart.

predictions about the physiological outcome. What these tools really help predict is 'what physiological outcome the patient can achieve if maximized care is continued'. In short, they predict physiological futility alone.

As we have seen in chapter 2, better predictions about the attainable physiological outcome alone may not always be enough to make good decisions. Correctly identifying the eventual neurological outcome of a particular patient does not guarantee correctly identifying the way the patient will value that outcome. Yet, while decision-makers struggle to delineate normatively futile outcomes, there seems to be an expectation that normative judgments are more or less determined and need only be sought out and identified. We see this both on the general and individual level.

On the one hand, the current endeavors to improve outcome predictions already have general normative futility judgments built into them. Physiological outcomes which all fall under CPC categories 1, meaning full recovery, to 5, meaning death (see Table 3.2³) are dichotomized in 'good' and 'poor' outcomes for both prognostic research and clinical practice. This would imply that normative futility judgments are clearcut and universal and that they can just be assumed. On the other hand, in many practices, decision-makers try to take individual preferences into account. While this would imply that normative futility judgments are *not* universal or clearcut, the assumption that these individual preferences, even if still unknown, are somehow already determined remains. The assumption is that these set personal valuations must merely be found out.

Both of these assumptions are problematic. First, as previously mentioned, even the dichotomized classification of CPC or GOS categories (see Table 3.2) has proved unstable. Before 2006, the category of severe cerebral impairment (CPC3) was considered a good outcome, while the majority of studies conducted after 2006 classified it as a poor outcome (Sandroni & Nolan 2015, Sandroni et al. 2018). Second, as we shall discuss in this paper, people may not be able to accurately identify whether a specific physiological outcome is acceptable or not, *especially* regarding outcomes that fall under CPC3.

CPC3 is the most controversial category because outcomes under CPC3 are least extreme, most ambiguous, and most diverse, including variations of severe motoric and cognitive damage and combinations of both. The common denominator is that these outcomes (unlike those that fall under CPC2 or moderate disability) typically do not allow for independent living and thus require extensive assistance.

³ Next to the Cerebral Performance Categories (CPC), outcomes are categorized under the Glasgow Coma Scale (GOS). While the categories themselves are the same, the numbers of the categories are inverted.

Patients with CPC3 are therefore dependent on others for activities of daily living. While this is currently categorized as a poor outcome, we shall see that normative evaluations of outcomes that fall under CPC3 are the ones most affected by normative uncertainty. Yet, because CPC3 is now grouped under ‘poor outcome’, deliberation about WLST emerges in practice for patients with such prognosis.

CPC	Description	GOS
1	Good cerebral performance / Good recovery	5
2	Moderate (cerebral) disability	4
3	Severe (cerebral) disability	3
4	Comatose or (persistent) vegetative state	2
5	Death	1

Table 3.2 – Glasgow Outcome Scale (GOS) scores and Glasgow - Pittsburgh cerebral performance categories (CPC)

In what follows, I show that the information necessary for normative futility and utility judgments, whether as universal standards or as individual evaluations, is much less stable than is assumed in neuroprognostic practice and research.

4. EVALUATING NORMATIVE FUTILITY

While it is unclear to what extent and in what way normative evaluation *should* inform decisions of WLST, it is primordial to identify and recognize precisely the uncertainty that normative futility judgments are subject to.

Whether or not continuation of treatment is considered normatively futile can vary depending on what outcomes are valued. Yet, the variation is not simply a matter of disagreement about values. By describing the difficulties our respondents experience when aiming to clearly identify ‘an acceptable and unacceptable quality of life’ for themselves or their next of kin, the following sections show that the evaluation of a particular physiological outcome is (a) personal and subject to change (§4.1), but also (b) dependent on expectations and gradual adaptation to new limits and possibilities (§4.2), and (c) dependent on a variety of intrinsic capabilities, external

factors, and the interaction between them (§4.3). The ways in which these different aspects affect one another are increasingly complex, especially considering the vast amount of new technological developments.

Ironically, with increased technological advances there is almost no limit to what bodily functions *can* be sustained, leaving the question open under what conditions bodily functions *should* be sustained. To be clear, in no way do I aim to answer this question here. Rather, by taking a closer look at how these questions are answered in practice, I aim to show to what extent answers may not be readily available to decision-makers and why this is so.

To remind you of our respondents in the different cases, I reproduced the table from the previous chapter below.

Cases	Reinhart	Willems	Vanderbilt	Felts
Relation to patient (Respondent)	Partner (Irene)	Partner (Judith)	Partner (Johanna)	Partner (Anke)
	Eldest child (Lukas)	Patient (Gerrit)	Eldest child (Natan)	Patient (Anton)
	Youngest child (Jason)	Neurologist (dr Bos)	Youngest child (Roos)	Only child (Jolien)
	IC Nurse (Alice)	Intensivist (dr Berger)	Intensivist (dr Harte)	Neurologist (dr Helmers)
	Intensivist (dr Valk)	Physiatrist (dr Raven)	Physiatrist (dr Naas)	Intensivist (dr Bent)

Table 3.3 – Respondents per case study (names were changed)

4.1. Views on quality of life and where to draw the line are diverse

In practice, the question ‘what treatment goal is normatively futile’, implies the question ‘what is an acceptable outcome for the patient’. Due to their normative nature, these questions are difficult to answer. Again, I do not attempt to provide answers to these questions. This chapter offers a reflection on the kind of considerations that relatives and clinicians involved in practical deliberation of such questions bring to the fore, focusing in particular on the complications and challenges they experience. Further philosophical inquiry may be necessary to evaluate the considerations weighed in practice and whether they indeed are ‘good’ reasons for withdrawing or continuing treatment for each specific patient.

'Where's the limit? That's a difficult question.' During my fieldwork, and contrary to what the question posed by dr. Valk suggests, it quickly became clear that the answer to what is acceptable or unacceptable for a particular patient is never a single point on the continuum. There is no clear, pre-existing boundary between the two; no 'limit' as such. Observations of conversations between clinicians and family members strongly suggest that the line has to be drawn for each case anew. Yet, during interviews, it became evident that, even on an individual level, finding a demarcation point is extremely difficult.

First, respondents almost immediately acknowledge that the boundary can be different for different people. This is especially problematic when the decision to WLST is made on behalf of another person. Judith, Willems' wife, emphasizes the differences between her and her husband. She refers to him as a 'do-person', always active, always busy doing something physical. Allowing him to become fully dependent on others, completely paralyzed and bedridden, even if fully aware, would reduce his quality of life tremendously. Judith phrased it as a 'punishment'. She, on the other hand, reads a lot, works with the computer, and studying is her hobby. Although she wouldn't wish that kind of dependency for herself either, Judith thinks her static activities would make it easier for her to live with severe physical disability.

Evidently, an identical condition may be experienced very differently from person to person, but even for a person who has the worst experience, that may not be sufficient reason for WLST. Even on an individual level, our respondents typically struggle to describe the normative line under which they themselves would no longer want treatment. In the Netherlands, these issues are publicly debated and such debates inform policy. In the past decennia, an overwhelming majority of Dutch citizens pronounced that loss of consciousness (Health Council 1994, de Beaufort 2005) and cognitive damage like the inability to communicate with (or recognize) loved ones is beyond what they would be able (or desire) to live with (Royal Dutch Medical Association, 1997).

Our respondents reflect this position well. Judith says:

He doesn't want to live like a vegetable. That was instantly clear to all of us [both me and his children], a vegetable is not an option. Remaining on a ventilator or being driven around like a zombie in a wheelchair, that's not an option. He didn't want that for himself, but he didn't want to do that to us either. Then we would've had to say, well, 'pull the plug', no matter how painful that is. We have discussed this a long time ago. His children know it too. I've known his wishes for years, and he's known mine. My children know mine. This is a steady fact for us.

Thankfully, Mr. Willems survived his cardiac arrest with excellent neurological function and his family did not need to make that decision at the time. Two years later, however, he was again hospitalized for a different health problem and, this time, his wife and children did indeed decide to let him go because his health had already deteriorated past what he considered desirable.

Before he passed away, however, Mr. Willems explained during our interview that although it is true he would have a harder time adapting to physical disability than his wife, to him, expressing himself was most important. Even if he could still think, remember and be aware, if he had no way of telling people about it - whether through speech, writing, or sign language - he wouldn't want to go on living. He said that his idea of humans was that we are social creatures and we need to be able to interact. When offered the possibility to interact through aiding technology or via a computer he remained resistant and much would depend on what that technology could do. In the end, he laughed and explained that maybe he was too old to adapt to these new tricks.

Most other respondents share the valuation of mind over body, especially when it touches identity. Felts' daughter, Jolien, and Lukas, Reinhart's oldest son, both, respectively, share their views on the loss of consciousness and self-expression:

I think the most important thing is whether or not you can talk and think well. Because if you can still talk, but you can hardly think, like people with advanced Alzheimer's, for instance, when you can't be consciously present in the world, I wouldn't want that. To remain physically, but without consciousness, that wouldn't be living for me.

You can take off all the body parts you want, I don't care. But if I can't express myself, by saying how I see someone, what I feel and think, then I'd rather you put me six feet under. And my father was just like that. ... We talked about it with dad too. He was always very clear. *This and this* I absolutely do not want. And you ought to respect that.

Vanderbilt's daughter, Roos, when asked where she would draw the line, considers the fact that a loss of mental capacity is not just detrimental to the patient but to those around them and that this would affect her decision too.

You're designated to one another. If I need to be taken care of, and the other person can't do it, I'd be more inclined to say, it's okay to end it here. If physically, or especially mentally, you don't know much anymore or you can't grasp what is happening around you, then that is really hard on the loved ones that surround you.

Roos' mother adds that not knowing what is going on inside her husband, not knowing how he experiences the world or his own life, is especially painful. 'If that part is missing, if communication is no longer possible, then it becomes really hard, for me and the children.' Just like Judith, the Vanderbilts point to limitations in the environment.

All four families implied that, when considering the patient's wishes, it wasn't just about the patient's wellbeing. There seems to be a limit to what the patient would find acceptable for his loved ones too. Their testimonies implied that there is a strong interaction between the patient's wellbeing and that of the family.

The Dutch physicians I spoke to voiced similar perspectives. Due to their professional experience, they are more often confronted with undesirable outcomes and see them from an emotionally more removed - indeed clinical - perspective. Reinhart's intensivist explains his priority.

I don't want to create vegetables. That matters to me most. [...] It's just not a quality of life that I ever wish for anyone. I know this is really subjective but I don't find this, what I would call a dignified and functional life, that's my personal view and I know others may think differently. Where is the limit? That's a difficult question. But if you're laying in a nursing home in a diaper as an adult and you can't do anything, you can't talk, or communicate in any way, I don't find that a life worth living. That is my conviction.

It is important to note that these citizen and professional views and reflections have a typically Dutch ring to them. In countries like Israel, WLST to avoid even prolonged disorders of consciousness are illegal. In the United States, where autonomy is the defining guide, decisions are made based on informed consent and informed refusal (Luce & Alpers, 2001) and, as was confirmed by my observations and conversations with physicians there, much is left to where the family draws the line. Yet, beyond the loss of consciousness and identity, that line becomes a lot harder to draw⁴. Even in the Netherlands, when one scratches the surface, driving questions into more detail or shifting the questions from their own person to their loved ones, or vice versa, previous assertions are retracted, contradicted, or replaced with general confusion and an inability to decide in the face of such overwhelming responsibility.

⁴ Some might argue that this challenge may indicate that the question for such cases should not be raised in the first place. However, outcomes that fall under CPC3 currently permit WLST in countries that allow WLST for futile outcomes. The question has thus become inevitable.

In essence, decision-makers are forced to make a kind of prediction about the value of a future scenario for a particular patient, often based on whatever information they have regarding the physiological prediction. We have now seen that this normative prediction is far from straightforward. While both the wellbeing of the patient and its impact on the next of kin is often taken into account, decision-makers are not necessarily in a position to predict the patient's limit. When given the room to deliberate about these considerations, decision-makers, especially next of kin, become hesitant to make such judgments.

4.2. Views on quality of life and where to draw the line are volatile

Central to this hesitation to predict normative futility or utility is the recognition that the 'limit' under which life no longer seems worth living is continuously shifting. We just saw that it differs from person to person and so depends on *whom* one imagines certain futures for. However, it also shifts for the *same* person, dependent on age and context, expectations, perspective, and prospects. In what follows, I describe the temporal dimension of futility judgments, or how quality of life is experienced as dynamic and highly dependent on the patient's ability and that of their environment to adapt.

Time and again, the patients and relatives interviewed testified how they gradually came to accept new boundaries. Remarkably, such gradual acceptance happens in either direction. Some respondents explain that you may want one thing at first but as soon as you have that, you want something more. Others say that, vice versa, you may have a limit in mind under which you no longer want treatment to be continued, but once the situation gets below that limit, it becomes the new baseline and you just do not want to lose more than that. Initial expectations, therefore, very much impacted the direction of the responses to a particular outcome.

Felts' doctors had significant reasons to believe that Mr. Felts' outcome looked grim. They informed the family accordingly, and although the patient woke up well, the family's lowered expectations were raised only very gradually. Felts' wife, Anke, recounts her own process of gradual acceptance.

Just the fact that he recognizes you, you go crazy from happiness. And every time you think one step further. We've got this far now, now we have even more hope. And not long after, he starts reading a plate on the wall that he could read from the bed 'the hospital is not responsible for loss or theft' some standard text. He suddenly started reading that to us. Very slowly, of course, and some words hardly worked but he started reading. Okay, so he can read. A couple days later, the occupational therapist, she said her number and he said 'Do you have a pen for me?' and he started writing down the number. I thought, okay, writing works.

Her daughter, Jolien, explains the psychological strain of fear that each new step brings, even when it is positive.

You're happy with what he can do every time, so that's very positive. Every time you think 'Oh! He can do this! Yes! He can read again.' But at the same time I was thinking, is this the end or will he go further than this? First, you're positive and then almost immediately you're negative, every time in that order. I thought, oh, he can do this. Then, will he be able to do this tomorrow?

In contrast, the Vanderbilts had received an uncertain, rather negative prognosis but they were convinced their loved one might recover if only he was given a chance. When Vanderbilt indeed started to regain consciousness, it looked to them like full recovery was no longer impossible. The process of continuously dropping down the scale of acceptable quality of life is well-explained by Vanderbilt's son, Natan:

In the beginning you think, [we want him to recover] one hundred percent, everything is still possible. Then there's a relapse, and you get smashed all the way back. Then you say, 'okay, but perhaps we can still make it to seventy percent'. And then you drop down again. And every time you change your scale, you weaken your expectations.

His sister Roos adds that whatever you end up with somehow normalizes.

One way or another, you start to think of it as 'normal'. It may be crazy, but for me, the way it is now, it just became very normal. You accept it, in a way. This is how it is and you can't change it.

And yet, it is hard to tell how robust these adaptations are. When Roos is asked what she would want for herself, given her experience, this gradual acceptance of severe states as 'normal' quickly crumbles.

Knowing where you are and what you're doing, if that's no longer possible. If I would no longer be aware of that, that would be my maximum, I think. But if I say it like that (starts crying) my father can't do that either. He can- he has passed- that maximum that I am now saying, he has long reached that, in the sense that, I don't know if he knows where he really is.

This scenario of gradual acceptance is very different from that of Willems and his wife who experienced a similar adaptation to new limitations in a way that seemed very natural to them, simply because they accepted it as part of ageing. Judith explains that once you accept the decline, it's only a matter of responding to it appropriately.

The moment I think, this [particular activity] is starting to go wrong, then we'll think about how to address the problem. We realize very well we are getting older. He just turned seventy. And well, you just can't do things the way you did them fifty years ago, or twenty years ago. So be it. We're very aware that with age things change, and they're different from what you expected them to be.

In the same line, one of the intensivists explained that when drafting an advanced directive, it's important to update it regularly, precisely because your situation might change so significantly that you may want to revise your preferred treatment. Someone's health may decline to the point at which they may want to sign a 'do not resuscitate'-order but their role and responsibilities in life may also change in ways that would affect their choices.

These shifts in valuation in relation to shifting expectations happen on the doctor's end too. Here, the doctor's specialty and their placement in the chain of care seem to be at play. From conversations with physicians who work with patients in long term care facilities, it appeared that their notion of 'good' and 'poor' outcomes differed significantly from those assumed in critical care. Dr. Naas, one of the psychiatrists says:

When I used to work at the hospital, then you'd get to cardiology and the cardiologist would say 'full recovery! You're doing really well' and then we'd come and have to tell the family the bad news that their loved one didn't even know how to get dressed, but the patient was easily walking around on the floor so no one could tell that there was serious damage. But he doesn't know where he is, doesn't know what time it is, doesn't know how to get dressed, doesn't know the inside of his own house or how to get from one room to another and these are the invisible things ... [a patient] might be able to walk again but aside from that he doesn't know anything anymore. And those are the people we see. The cardiologist is pleased with the outcome, the neurologist is very pleased, and then the practical reality is dramatic.

The meaning of good and poor outcome can thus vary along the chain of care too. Depending on a professional's position in the process, their boundaries seem to change.

When the experience of 'good' or 'bad' outcomes depends on expectations and the ability to accept new situations, better outcomes may depend on flexibility, resilience, and a supporting environment. Judith explains the ways she and her husband had to continuously find new ways of adapting to increasing loss of short term memory and physical ability.

Once you can't do something anymore, you've got to learn to live with that. He systematically forgets to turn off the coffee machine, so now I turn it off. He systematically puts the fire too high, making the pans boil over, so I turn it lower and then they don't. If I don't put down a note for groceries, he forgets to do them. So, I put down a note. We've found a new *modus operandi* ... We can't go on long walks anymore, but cycling still works, so now we cycle. But the bikes are 25 kilos a piece. Putting them on the camper became too heavy for him. Well, then we build a winch and pull them up automatically ... Time and again, we've found ways to compensate for lost capacity and ability.

While these are simple examples for compensating moderate neurological damage, they show how the quality of life actually experienced is very dependent on how diminished quality is compensated for, by other people or by environmental adaptations or both. Such understanding of quality of life further complicates an anticipatory evaluation of what will or will not be acceptable, especially when new technological developments allow compensation for severe disability. Increasingly, acceptability seems to depend on the flexibility of the patient and their environment (Roulstone 2016). The latter is also dependent on what kind of compensation is offered at an institutional level, as many scholars working on disability and technology have argued as well (National Academies of Sciences, Engineering, and Medicine 2017). Typical examples are whether or not public places are made accessible for wheelchairs and whether or not health care covers artificial limbs for amputees and enhancement technologies or transplants for visually impaired people. For neurological damage, it may be harder to predict what support and compensation is available, especially if one doesn't know whether the damage will be cognitive or motoric. Moreover, for patients with prolonged disorders of consciousness, the question remains whether compensation for lost ability, if possible, is *always* the best way to go.

In any case, the notion of normative futility is further complicated by the initial expectations, the fluctuations in valuation, the flexibility and resilience of the patient, and the realization that the lives of patients can be significantly improved by adapting the environment to the challenges their limitations cause. At the same time, this insight also reveals room for improvement. As we shall see in the next section, family members who became aware of this dynamic indicated the need for improved coordination of care to better support patients.

4.3. Views on quality of life and where to draw the line are context dependent

Whether further treatment is considered normatively futile depends on what quality of life can be achieved. In the previous section, we saw that the capacity of people to gradually adapt to new limits, combined with the observation that experience of

limitations depends on the support offered by the environment, can impact what physiological outcomes are seen as acceptable. This has a determinant effect on whether continuation of treatment is considered normatively futile in practice. As such, these considerations raise the question whether more can be done to improve support, thereby increasing quality of life, and making fewer treatments futile.

At first, Jolien quickly dismissed complete paralysis as an acceptable outcome for herself, even if she were still able to think and talk clearly. But then she remembered a boy who could move his fingers and head and with that, he could control a car. Suddenly, losing control over the rest of her body didn't seem so bad anymore. While this is a simplified, imagined example, cases of such shifts happen in reality and it shows that something like gaining mobility on a scenario that didn't allow independent movement can make all the difference.

In light of the difference that technological mediation can offer, quality of life depends to a degree on what our healthcare services and our social and material environment can provide. The available support can make or break the kind of prospects someone has.

From his son's perspective, Vanderbilt did not get the kind of support that could have maximized his recovery. Natan says:

My father didn't receive the care he needed... If that care existed, if people understood his problem, could offer the tools, the budget, the living situation or nursing, then I would say, go for the maximum. Go for it. But, you know, the whole cycle that follows is not good enough to bring each situation to a good end. That's why I would say, trust the doctor, take their advice to heart... They can influence the system. Hopefully, they know what happens after their part too. They may not know what the home care will be like, but up to a certain point, until rehabilitation, they should know what's at play. And I think, if I had known that this is how it would be today, I would've said, stop at the ICU already. Because the continuation of care, the lack of coordination, and the problems that it caused, resulted in a quality of life that is zero.

Three considerations are especially important to note. First, good and clear communication about what people can expect from long term care is crucial in deciding whether or not to prolong care. Second, it raises the question whether providing better long term care would lead to other choices regarding WLST. Finally, as much as Natan would like to trust that doctors know what happens on an institutional level beyond their specialty, our fieldwork shows that precisely the lack of coordination and exchange makes this kind of foresight very difficult.

Coordination along the chain of care is reportedly lacking in the Netherlands and new initiatives have been set up in response (Zadoks & Gijzen 2019). My conversations with physicians working in intensive care show that they do not, in fact, know very well what happens after patients leave the ICU. They admit that, currently, knowledge transfer between the different stages of care is suboptimal, on the individual case-to-case basis, as well as on the more general, scientific level. The physicians as well as the patients and the families I spoke to experienced how poor coordination between the different stages of the care continuum has an effect on continuity of care. As a result, initiatives to invite patients for follow up meetings are increasingly being proposed, and they are received with enthusiasm. Although the clinicians I spoke to found insights about the trajectory further down to be extremely relevant towards informing the family about long term options during the acute phase, they felt these insights were still insufficiently communicated back to them and those who would be served most by them.

Sometimes, these insights are about *which* therapeutic goals are valuable. Here too, misalignment about the value of a specific treatment can greatly affect the patient's quality of life. For instance, therapies that do not result in improvement are ended because they are regarded as futile. But when therapies are taken away from these patients, their abilities often diminish quickly. For Vanderbilt's wife, this was one of the most painful experiences in the care given to her husband. Even though the issue has been partially remedied since, Johanna still gets very angry and upset when she recounts what happened.

Talking is becoming increasingly harder. Speech therapy was stopped. Music therapy was stopped. Yeah, 'because there's no progress'. I got so angry then! On what ground! (bangs her hand on the table, then cries) His therapies are canceled because nothing sticks—how do you call that—if you learn something today and pick it up tomorrow, in a normal learning process. But that's not applicable to him! He can't pick up where he left off. He has to start again from the beginning! But he *needs* that *basis* to be able to still express *something* and then it just gets--then--it's not allowed anymore, financially. Luckily, music therapy has now been restarted for a bit, after I insisted very much, that he gets happier from music, that on those days, he talks, he is more vocal, everything is nicer.

Clearly, family will, once aware of the way adapted care can impact their relative's quality of life, actively seek out more appropriate care. By informing next of kin early on about the impact of available care, they may be better equipped to make good decisions or, at the very least, request information about the options. Conversations about existing options and their effects or lack thereof can furthermore lead to improvements on the institutional level.

Generally speaking, if the decision to withdraw life-sustaining treatment is based on what quality of life is acceptable, and the latter depends on what the care chain offers further down the road, it is worthwhile to investigate if there may be alternatives to WLST and for whom these alternatives would be better.

Summing up, section 4 lays out three ways in which uncertainty complicates straightforward identification of normative futility. I argued that evaluations of what is (un)acceptable quality of life in practice, aside from being diverse, depend on a variety of considerations and factors, including the person's own resilience, flexibility, and ability to adapt to new limitations, as well as that of the patient's immediate support system. Furthermore, the support that can be offered at an institutional level has a great impact on the patient's ability and that of their family to respond to limitations. This means that normative futility judgments are quite undetermined and understanding what the patient's underlying values are and attempting to directly translate those to a judgment of acceptable quality of life may not be an accurate way to assess whether further treatment is futile. In identifying the expected quality of life, a balance has to be found each time between value judgments and the contextual circumstances. At the very least, dichotomizing the value of outcomes in 'good' and 'poor' purely based on physiological outcome alone seems rash, especially for outcomes that are quite undetermined such as the range of outcomes under CPC3 that are currently evaluated as poor. With so many variables impacting the quality of life that the patient will ultimately experience, uncertainty about what outcomes are to be considered as acceptable abounds. This further complicates a straightforward notion of futility and the decision whether or not to withdraw life-sustaining treatment.

5. PULLING APART PHYSIOLOGICAL AND NORMATIVE UNCERTAINTY

Empirical investigation indicates that normative futility judgments are not as distinct nor as stable as current neuroprognostic practice assumes. They are in fact subjected to a lot of uncertainty and not just due to uncertainty about what physiological outcome can be achieved. As shown above, the kind of quality of life a patient will experience and whether that is acceptable to them depends on more than physiological factors alone. Therefore, even if accurate prediction of neurological outcome following continuation of treatment *could* be achieved, uncertainty may remain about what outcomes ought to be prevented or pursued. I argue that this uncertainty combined with the uncertainty regarding the physiological outcome, together form the overall neuroprognostic uncertainty that decision-makers struggle with. If we *must* define neuroprognostic uncertainty as a single concept, it should at the very least incorporate these two aspects (e.g. the definition in table 3.4).

Neuroprognostic uncertainty	<i>Uncertainty whether the proposed treatment can (is likely to) physiologically achieve a treatment goal that (likely) constitutes an acceptable outcome</i>
------------------------------------	---

Table 3.4 - Definition of neuroprognostic uncertainty

Yet, such a definition is not very helpful. To be able to resolve neuroprognostic uncertainty, it is essential we distinguish between the two uncertainties, understand the challenges either one poses, and address each of those challenges. That said, it is understandable that the two uncertainties have been confounded for so long. They are strongly interconnected. Normative futility judgments still help inform physiological futility judgments and normative futility judgments require information about the physiological outcome. But neither depends *solely* on the other. In what follows, I pull apart and clearly define the two types of uncertainty at stake, mirroring the two types of futility judgments.

5.1. Two distinct uncertainties requiring two distinct predictions

When deciding whether or not WLST for a patient in postanoxic coma is justified, decision-makers face uncertainty about the neurological outcome of the patient (*physiological uncertainty*), about the acceptability of that outcome (*normative uncertainty*), or both. While physiological uncertainty is caused by imperfect (or lacking) knowledge of a treatment's efficacy, normative uncertainty is due to imperfect (or lacking) knowledge about which treatment goal is worth pursuing.

Physiological uncertainty	<i>Uncertainty whether the proposed treatment can physiologically achieve the (foreseen) desired effect</i>
Normative uncertainty	<i>Uncertainty whether the (foreseen) physiological effects of the treatment constitute an acceptable outcome</i>

Table 3.5 - Definitions of physiological and normative uncertainty

The inclusion of 'foreseen' effects in the definitions above (Table 3.5) clearly show how uncertainty about the one can add more uncertainty to the other and vice versa. Clearly, pulling them apart is no easy task. However—while we should keep their relation in mind—it is vital that we also study the two in isolation because they do not *solely* depend on one another. When there is no remaining physiological uncertainty about what outcome can be achieved, normative uncertainty can still remain. Vice versa, when there is no more normative uncertainty about what outcome is acceptable, physiological uncertainty can still remain. The two are simultaneously independent *and* interdependent.

As we have seen, judgments of normative futility are subject to so much uncertainty, that the normative futility of continued treatment may not yet be determined. Because it is so undetermined, it is more accurate to say that—rather

than *identifying* the normative futility or utility of continued treatment on the basis of pre-existing values—one is obligated to *predict* it. This means that neuroprognostication requires two—not one—predictions.

One could argue that, with so much uncertainty, having standard categories for normative evaluation is still the most workable solution for practice, even if it is imperfect. Indeed, I would argue that the need for practical solutions necessitates such imperfect strategies until they can be improved. Based on my analysis, however, the current practice can and *should* be improved. Currently, outcomes that are predicted to fall under CPC3, in countries that allow WLST, give reason to withdraw treatment and in some or even many of these cases that may well be the best response. However, with the acceptability of such outcomes being so dependent on external context and environmental and institutional support, policy makers should at least discuss whether there are alternatives to WLST for these patients and whether they should be covered and supported on an institutional level. In the Netherlands, we already see some recent efforts being made in that direction⁵ Furthermore, in the final two chapters we will see that the predictions made about whether certain outcomes are acceptable or not can be self-fulfilling and thus make it more likely that they will or won't be. This forms an additional reason to be extra careful when predicting normative futility.

5.2. Who predicts what?

Pulling apart physiological and normative uncertainty has an additional benefit. While both the physiological and normative questions are difficult to answer, it isn't always clear *who* should be answering *what* question either. Based on their expertise, it may seem logical that physicians decide on the physiological answer and family on the normative one. Yet, that is not necessarily what happens in practice and this is largely due to the confounding of the two issues.

The ways in which physicians and family collaborate or negotiate also varies greatly, depending on the country and cultural context. As seen in the previous chapter, medical professionals in the Netherlands often try to obtain information about the patient's wishes or what the family wants and aim to get on the same page, while also keeping in mind the national guidelines regarding futile treatment. The entanglement of physiological and normative uncertainty makes decisions especially difficult. And to the extent that normative futility judgments are dependent on physiological futility judgments, this is understandable. Even when there seems to be a clear indication of preferences, decisions are ultimately made based on whatever limited knowledge is available about the physiological question regarding the patient's chances. If that remains unknown, the risk is inevitable.

⁵ I discuss these efforts in more detail in chapter 7.

Dr. Valk, intensivist:

What you often hear is, 'he would have never wanted that kind of dependency, having to live in a nursing home', in which case I am inclined to suggest to stop treatment, but I find that very hard to do, because I may stop treatment of someone who may wake up okay. On the other hand, I think, well, I also protect several people from waking up horribly. But yeah, that is... that is really difficult.

Felts' neurologist, Dr. Helmers, illustrates with a recent example:

We recently had a very young patient, a woman of around thirty with severe encephalopathy of whom the family indicated that 'were she to come out seriously damaged or disturbed, we would rather spare her that'. On the other hand, she was 29 or 30 and she was improving. We all knew that the chances were high that she would keep disorders but we knew that there was also a large potential for improvement. And due to her age and that potential for improvement, we couldn't give up, we couldn't. While we did know, the chance this woman recovers in ways that the parents don't actually want, is really high. Yes, and that is, of course, an enormously difficult situation then.

This example shows that family expresses, as best as they can, the kind of quality of life they think the patient would find acceptable which reduces normative uncertainty. Yet, even when the clinicians take this to heart, the physiological uncertainty gets in the way of easy decision-making. And as long as there's no reliable answer to what the patient's outcome will be, our respondents say that families often indicate that they want doctors to decide, not always because they want to be relieved of the burden it puts on them but often also because they're afraid they will make a mistake. They want an acceptable outcome for the patient and avoid the worst case but may not feel they are in the best position to make the call.

Vanderbilt's son, Natan:

Because you want the doctors to decide. In a way, that's what you want. You want them to spare you the very worst. And as a family, you want, you try to create a situation that the patient would find acceptable as well.

Dr. Helmers experiences the same concern when talking to families in the Netherlands:

It often happens in conversation that the family is enormously doubtful, naturally, and that you can indicate, 'yes, but listen for a moment, you don't

actually have to take that decision. We have to make that decision.' And most families find that really agreeable. When you take that away from them. Most appreciate it when you say: you don't have to decide, I decide.

When new prognostic tools provide better predictions about the outcome, however, thereby reducing physiological uncertainty, this could allow more room for decision-making based on proper predictions of normative futility and utility judgments. Regardless of who is ultimately best equipped to make these normative predictions, the need for such predictions and the challenges they entail must first be acknowledged as such.

Clinical practice shows that there are two separate uncertainties that need to be considered and that different parties may be better equipped to answer different questions. Making the distinction between the two uncertainties explicit, can thus furthermore help decision-makers in their deliberations and the allocation of responsibilities.

6. CONCLUSION

Decisions to withdraw life-sustaining treatment are informed by medical futility (and utility) judgments. However, it is well-established that neuroprognostic uncertainty complicates such judgments. Indeed, technological innovations are researched and developed to reduce neuroprognostic uncertainty. In order to ensure responsible innovation of neuroprognostic practice, we must first better understand the uncertainty at stake. In this chapter, I argue that the uncertainty that is sought to be reduced in neuroprognostication is an entanglement of two distinct types of uncertainty—physiological and normative—mirroring the two different types of futility.

While reducing the uncertainty with regard to physiological futility judgments has been central in neuroprognostic innovation, the uncertainty decision-makers experience regarding normative futility judgments has been grossly overlooked. Instead, the ways in which neuroprognostic tools are investigated, assumes that there is *no* normative uncertainty. This is because physiological outcomes in prognostic research and scientific publications are already tied to normative judgments in ways that imply certainty about which outcomes are good and which ones are poor. To offer more balance, the empirical material used in this chapter focused on better understanding the challenges when trying to make judgments of normative futility.

In conclusion, developing better prognostic tools to more reliably predict the physiological outcome of the patient is crucial in reducing neuroprognostic

uncertainty. But it is not enough. In addition, more in-depth investigation is required on how to better predict the normative futility of a specific outcome. These predictions must furthermore be understood as dependent on external factors, some of which being within our control. Since choices can be made, one must consider which choices are best. Importantly, prognostic tools cannot—and thus should not—include normative futility judgments. Neuroprognostic tools should refer strictly to physiological outcome categories. Further classifying these categories into normative ones should be a separate step. It is furthermore likely that outcomes under CPC3 cannot currently accurately be dichotomized as either good nor poor but require their own, less determined classification. Even if standard categories for normative evaluation is the best one can do, those categories should not be inherently embedded in the prognostic technology for predicting physiological futility. This is not just essential for deliberations about WLST, it also greatly affects the research of neuroprognostic tools as their sensitivity is calibrated not on the different physiological outcomes (in the form of CPC categories) but on the normative categories (good and poor).

A necessary step to more responsible neuroprognostic innovation is thus to clearly separate between predictions about what outcome can be achieved and predictions about what outcome is acceptable. As such, these two different predictions can address the two distinct uncertainties at hand. Finally, because physiological and normative issues are taken as one, it's hard to differentiate between the two and allocate responsibilities and decision-making power to the appropriate parties. Since both fall under general notions as 'medical futility' and generic 'uncertainty' about the outcome, it is difficult for different decision-makers to bring their specific expertise to the table. For research, clinical practice, and deliberation, it is essential physiological and normative issues are distinguished and addressed separately.

Chapter 4

Chasing Certainty After Cardiac Arrest: towards responsible implementation of cEEG for neuroprognostication

ABSTRACT

Recent research shows that continuous EEG-monitoring (cEEG) can substantially improve prognostication of outcome for patients in coma after cardiac arrest. This raises expectations that the use of this technology will make it easier to decide whether or not to withdraw life-sustaining treatment. In this chapter, I investigate the plausibility of that expectation by exploring what effects cEEG has when it becomes part of a socio-technical network in an Intensive Care Unit (ICU). The analysis focuses on (1) the way patient outcomes are constructed, (2) the kind of decision support these outcomes provide, and (3) how cEEG affects communication between professionals and relatives. I argue that cEEG in some cases can indeed eliminate or reduce uncertainty but that it also raises new concerns. Moreover, I point out how the actual impacts of cEEG in practice hinge on *how* cEEG is designed and implemented in practice. I end with an inventory of issues that should be addressed to ensure responsible further development and implementation of cEEG.

Chasing Certainty After Cardiac Arrest:

towards responsible implementation of cEEG for neuroprognostication

1. INTRODUCTION

On Saturday, October 26th, 2013, my daughter and I left the house. My husband, Chris, had been alone since 9 am. We had almost reached home when Chris called me in distress 'where are you? I need you'. I immediately had this feeling; this isn't right, this *really* ain't right. We raced home as I said 'Hold on. We'll be home in three minutes.' Meanwhile, my daughter called 911. As we jumped out of the car and into the house, we saw Natan, my son, standing on the driveway by chance. He had no idea. Chris was still sitting at the table when we ran in but as we reached him, he collapsed. Natan and I started reanimation but Chris had to vomit and his head was quickly turning blue. I got really scared that he wasn't getting the oxygen we tried to give him through mouth-to-mouth breathing. It took the ambulance seven minutes to arrive. Reanimation took a long time and after seven defibrillations and lots of medication, Chris was stable enough to go to the hospital. But even on the way to the hospital he had another cardiac arrest and more defibrillation. (Vanderbilt's wife)

Chris Vanderbilt remained in intensive care for weeks. Prognostic tools, although limited by a remaining degree of uncertainty, indicated a poor outcome. Nevertheless, the Vanderbilts managed to convince the physicians to keep going by pointing to video footage that demonstrated signs of consciousness, indicating that withdrawing life-sustaining treatment might arguably strip Chris of a viable chance. Even so, both family and treating physicians were aware of the danger continuation of care might hold for the patient and their quality of life; Chris could survive in ways he might consider worse than death.

The (im)permissibility of withdrawal or withholding of life sustaining treatment (Rydvall & Lynöe 2008, Welie & Ten Have 2014, Ko & Blinderman 2015), as well as the difference between the two (Thomson 1976, Husak 1979, Tooley 1980, Cartwright 1996, Rachels 2001, Foot 2002, Levin & Sprung 2005), have been extensively discussed in the philosophical literature. Since many countries legalized withholding and withdrawal of life-sustaining treatment for irreversible cases, however, the ethical issues in those countries shifted to the level of medical practice, where they reappear in a more practical form. Medical professionals as well as relatives of patients in coma after cardiac arrest now have to answer the question 'Should we continue treatment for this particular patient or not?'

As seen in the previous chapter, any answer to this practical question has to address two sub questions. First, one needs to know the minimum quality of life that should

be achieved to make continued living worthwhile. In the previous chapters, we've seen that, in practice, this normative criterion is usually inferred from what is known about a specific patient's wishes and values but it is also dependent on cultural and political context. Second, information is needed about the chances that a particular patient will be able to attain that minimum level – that is, about a patient's prognosis. As seen in the example above, my first exploration of current prognostic practices for patients in coma after cardiac arrest shows that if a patient's prognosis is uncertain or lacking, the option to discontinue care has created a moral dilemma. Medical practitioners and families are forced to choose between two potential harms: continuing care with the risk that the patient will survive in ways they would have never agreed to, or withdrawing care with the risk of allowing the patient to die, never knowing if they could have recovered.

It need not surprise, then, that there is a strong urge to reduce the uncertainty underlying this dilemma by developing and further improving prognostic tools. In the case of prognostication of patients in post-anoxic coma, continuous electroencephalogram (cEEG) monitoring was recently shown to reduce prognostic uncertainty. It not only provides reliable predictions of poor outcome, but also enables a robust predictor of good outcome (Tjepkema et al. 2015, Hofmeijer et al. 2015, Spaletti et al. 2016, Ruijter et al. 2019). The promise of increased certainty also raises expectations regarding care. If results are more certain, they might (1) provide more clear-cut information for medical professionals, (2) offer better guidance for decision-making, and (3) be easier to communicate to relatives.

Obviously, a reduction of prognostic uncertainty is desirable, especially given the high stakes. Moreover, it might indeed positively impact professional understanding, decision-making processes, and communication. However, past research from philosophy of technology and science and technology studies has shown that new technology usually does both much more than, as well as things different from, what it is intended to do (Geels & Smit 2000, Parthasarathy 2005, Akrich 1992, Verbeek 2005, 2008). These broader impacts, moreover, can considerably influence the societal and ethical acceptability of a new technology. I therefore

- (1) explore the broader, unintended effects of cEEG-monitoring on care practices in intensive care units (ICUs), and
- (2) make an inventory of societal and ethical issues that need to be addressed to ensure responsible development and implementation of cEEG.

I first introduce current practices of prognosticating for patients in coma after cardiac arrest and the promise of cEEG monitoring (§2) and then provide more background regarding the theoretical framework (§3) and methods (§4) I used. I subsequently discuss my findings, which roughly cluster around three themes: the construction of patient outcomes (§5.1); the decision support provided (§5.2) and

how to communicate results (§5.3). I show that in each case the increase in certainty that cEEG offers is replaced by new uncertainties. I end by discussing implications for the design and implementation of cEEG in the ICU (§6).

2. POST-ANOXIC COMA, PROGNOSTICATION AND CONTINUOUS EEG-MONITORING

For about nine in ten patients, cardiac arrest is lethal (Hassager et al, 2018, 989). Patients who are successfully resuscitated usually remain unconscious in *postanoxic coma*, meaning coma after oxygen deprivation. When these patients arrive at the ICU they are sedated, intubated, and their body temperature is often controlled to prevent fever and further brain damage (989). About half of these patients survive with favourable neurological function (989). Valid neurological prognostication is vitally important because patients who die after the first twenty-four hours, mostly die from withdrawal of life-support in anticipation of severe neurological damage (989, 994). Existing neuroprognostic indicators commonly include the pupillary reflex test, the corneal reflex test, and the Somato-Sensory Evoked Potential (SSEP) test¹. These must be performed at least 72 hours after the arrest (991). All three tests look for a response from the patient: changes in pupil width after a light stimulus, movement of the eyelid after a touch stimulus, and electrical activity in the brain after an electrical stimulus of the wrist, respectively. If such responses are bilaterally absent, recovery is highly unlikely. However, these tests only predict poor, not good outcome and only for up to half of patients with poor outcome (Zandbergen et al. 1998, 2006; Sandroni et al. 2018). In practice this means that in the majority of cases, relatives are told to ‘wait and see’ whether a patient will wake up.

Several studies have shown the added value of EEG-monitoring. Highly malignant EEG after 72 hours reliably predicts outcome for half of the patients with poor outcome (Westhall et al. 2016). Furthermore, the addition of cEEG enables an overall prediction for up to half of *all* patients, in some studies as soon as 12 to 24 hours after reanimation (Tjepkema-Cloostermans et al. 2015, Hofmeijer et al. 2015, Spaletti et al. 2016). The increase in predictive capacity is largely due to cEEG’s ability to predict good outcome. The cEEG-based test requires that non-invasive sensors are attached to the patient’s head, preferably as soon as the patient arrives at the ICU. These sensors register electrical activity in the brain. The registrations are stored and represented in the form of graphs on a monitor. This monitor may be at the bedside or available off-site.

¹ Other imaging technologies like CT and MRI as well as biomarkers (eg neuro-specific enolase (NSE), protein S-100B, and the axonal injury marker tau) can be used to complement the multimodal prognostic assessment but should not be used in isolation. (Hassager et al. 2018)

In order to predict good and poor outcomes, interpretation of the cEEG graphs requires that specific patterns are identified. This can be done in three ways:

- (1) qualitative: through visual observation by a doctor (typically a neurologist);
- (2) computer-assisted: using human extracted quantitative features to create manmade algorithms which translate the specific features into a prognostic score;
- (3) machine learning-based: (a) including traditional machine learning in which predictive models are created after explicit features were provided (b) including deep learning techniques in which a neural network is trained to independently predict without being fed explicit features.

The third option has only emerged very recently (Tjepkema-Cloostermans et al. 2017, 2019). The predictive value of cEEG slightly varies depending on the mode of interpretation. Recent publications suggest that computer-assisted interpretation is better at predicting outcomes than visual interpretation, while machine learning techniques may eventually outperform both (Tjepkema-Cloostermans et al. 2013, 2017, 2019).

In sum, cEEG-monitoring seems to have several advantages. Its addition to existing prognostic tools allows for:

- increased reliability of prognostication of poor outcome
- prognostication of good outcome
- a faster provision of the prognosis (at 12 or 24 hours, instead of at 72 hours)

Based on the evidence available, cEEG monitoring was recently considered sufficiently robust to be included in the Dutch clinical guidelines for prognostication of patients in post-anoxic coma (Horn et al. 2019). Other countries also increasingly include cEEG in their multimodal prognostication and, as of 2021, the new European guidelines for multimodal prognostication list EEG's prognostic values alongside those of corneal and pupillary reflexes, SSEP, NSE, and CT or MRI, requiring two or more of these indicating poor outcome to indicate that a poor outcome is likely (Nolan et al. 2021).

3. THEORETICAL BACKGROUND: TECHNOLOGY, ITS PERFORMATIVITY AND ARCHITECTURE

Inclusion of the technology in guidelines does not imply that seamless integration in practice will follow automatically. Technology assessment for the development of clinical guidelines tends to focus on a narrow set of outcome measures, linked to the technology's intended purpose (Boenink 2012). In the case of prognostic technologies, validity, reliability, and predictive value are considered most important in the decision to implement. As research in science and technology studies (STS)

has amply shown, however, technologies tend to do much more than intended. They are *performative* (Pickering 1995, Berg & Mol 1998, Law & Singleton 2000, MacKenzie et al. 2007, Licoppe 2010, Wallenburg et al. 2019): while aiming to measure aspects of the world, they also reconfigure that world. For instance, as I shall demonstrate in this chapter, when measuring the state of a patient's body, technologies also reconfigure the interaction between a patient and a doctor, the way health is experienced, and even the concept of health and disease itself.

Several concepts put forward in STS highlight these performative effects. First of all, Akrich pointed out that technological devices have a *script* that defines, among many other things, roles and responsibilities of users, what type of actor can be a user, and what values are important (1992). Moreover, these devices need to become embedded in social-material networks to fulfill their function. As Parthasarathy (2005, 2012) has shown, what a test actually is and does depends on the way it becomes embedded in such a network. Rather than asking how a test impacts medical and social practices, we should ask how a new testing modality and existing practices *configure each other* and lead to new *testing architectures*. Finally, answers to these questions are not always straightforwardly available when the more complex parts of the technology are *black boxed* and only input and output are visible (Latour 1987). We must therefore unpack, as much as possible, the ways in which inputs are turned into outputs.

When considering the desirability of new technologies, the point is that desirability does not depend only on the effectiveness with which intended aims are realized. Rather, overall impacts on clinical and social practices are crucial. To enable timely reflection on the social and ethical desirability of cEEG-monitoring, we must investigate the performative effects of cEEG in practice.

4. METHODOLOGY

As explained in chapter 1, investigating cEEG in practice is possible by looking specifically at ICUs that have already installed cEEG for research purposes. Although cEEG-based information should not be used for prognostication due to its experimental status, such *liminal innovation practices* (Mertens 2018) already show how the technology influences—and is itself influenced by—the sociomaterial networks it becomes part of. As mentioned in the thesis' introduction, I conducted ethnographic fieldwork in two Dutch hospitals and, in order to contextualize the findings in Dutch practice, one hospital in the United States. All three hospitals had introduced cEEG-monitoring for postanoxic patients for research, before it was included in guidelines for clinical decision-making. Over twenty patient cases were followed and analysed and four were selected for in-depth analysis. This particular

selection was based on the empirical insight that concerns greatly varied depending on the predicted outcome, and even more so in retrospect, when compared to the actual outcome. The four selected cases thus aimed to capture the widest possible diversity² in, and combinations of, predicted and real outcome (see table 4.1), as these variations offered the greatest variety of concerns.

Cases	Prediction based on existing practice	cEEG-based prediction	Actual outcome
Reinhart	Poor	Poor	Died after withdrawal of life support
Willems	Good	Good	Lived without severe neurological damage
Vanderbilt	Uncertain / optimistic	Poor	Lived with severe neurological damage
Felts	Uncertain / pessimistic	Poor	Lived without severe neurological damage

Table 4.1 – Predicted and actual outcome of the 4 case studies (*patient names were changed)

To limit contextual variety for comparative purposes, all four cases were taken from the Dutch settings. Since cEEG was part of the research practice, the standard care protocol was followed for decision-making regardless of the outcome predicted by cEEG. Where relevant, I expand the comparison with the US practice.

For each case, I held open, semi-structured interviews with the healthcare professionals involved, surviving patients (when possible), and family members. Twenty-five respondents were interviewed, including four intensivists, five neurologists, two rehabilitation doctors, one IC-nurse, one technician, two patients, four partners of patients, and five of their children. Several medical interviewees were also key contributors to the development of cEEG as a prognostic tool. I obtained informed consent from all participants. All interviews were transcribed, personal data were anonymized, and names changed. Approval was granted by the Ethics Committee of the University of Twente and approved by the individual hospitals' ethics boards.

All data were analysed in Atlas-ti, combining inductive and deductive coding. The starting point was the question: 'what are people concerned about?' Initially I coded inductively, allowing matters of concern to emerge from the data. Once I developed an initial list of concerns, I analysed mostly deductively. For this chapter, I focused primarily on concerns related to the new prognostic technology. Additionally, the

² My study contained only one case where cEEG predicted a good outcome, when the actual outcome was poor. However, the death of the patient was not due to neurological damage and therefore, it was not included for in-depth analysis.

observation reports and interviews comprised the basis for conceptual analysis of the technology's performativity. For valorisation purposes, I presented the findings to an expert panel and organized a multi-stakeholder workshop. Participants' responses were included in the data set.

5. CEEG IN PRACTICE: CHALLENGES AND UNINTENDED EFFECTS

I clustered my findings into three groups, each relating to expectations raised by the increased prognostic certainty cEEG offers:

- test results will offer more clarity about a patient's neurological state (5.1. facilitating interpretation)
- decision-making on individual cases will be easier (5.2. facilitating decision-making)
- communication with relatives will be easier (5.3. facilitating communication)

In all three respects, my observations of how cEEG works out in practice indicate that the technology does not fulfill these expectations in a straightforward way.

5.1 Facilitating interpretation

As indicated above, current prognostic tests for postanoxic coma have low predictive value. The evidence available shows that the addition of cEEG improves predictive value and may thus reduce uncertainty. When inquiring how the results of cEEG are produced in practice, however, it became clear that in the process many other uncertainties emerge that need to be addressed. Moreover, the resulting practices inevitably, yet invisibly, embed specific norms and values. In what follows, I will unpack both issues by reconstructing what is required to produce cEEG-based test results.

5.1.1 Identifying patterns

Prognostication based on cEEG-monitoring requires identifying patterns. The evolution of patterns over time will form the basis of the prediction. It is not obvious how these patterns should be distinguished, however. EEG merely registers electrical activity in the brain. The main variables researchers look at are amplitude and frequency of the electrical activity, as well as (ir)regularities. This process can lead to different categorizations of patterns. Table 4.2 summarizes categorizations used by the researchers in the practices I observed.

Willems' neurologist, also a cEEG researcher, explains:

I have to add that these categories, to an extent, they're arbitrary. This is just what *we've* come up with. It isn't so that these are *the* generally accepted categories in which you can classify the EEG for this patient group, but rather *a way* of categorizing.

He added that there is, for instance, disagreement whether or not to distinguish between GPD and epileptiform. The relevant categories have thus not been fully standardized yet. Other research groups, for instance in the US, Italy, and South Korea, use slightly different classifications (Spaletti et al, 2016; Oh et al, 2015). Accordingly, there is uncertainty regarding whether the patterns currently used are the most informative ones. There may also be other patterns that have not been identified yet.

Patterns	Amplitude	Dominant Frequency	Characteristics or other distinctions
Iso-Electric	/	/	No brain activity recorded (flatline)
Low voltage	< 10 microvolt		Amplitude of signal is low
Burst suppression	< 10 microvolt & bursts > 10 microvolt		Identical and non-identical bursts are above 10 microvolt, rest is suppressed
Diffusely Slowed	> 10 microvolt	< 8 Hertz	Normal amplitude, but frequency is lower than normal
Normal	> 10 microvolt	> 8 Hertz	Continuous activity, normal amplitude and frequency
Generalized Periodic Discharges (GPD)	Epileptic activity		Discharges occur simultaneously on all channels and at regular intervals
Epileptiform	Epileptic activity		Constant epileptic discharges

Table 4.2 – Categorization of cEEG-patterns to prognosticate patients in postanoxic coma - in Dutch scientific research

5.1.2 Classifying EEG-measurements

Interpreting the cEEG-measurements also poses challenges. Traditionally, EEGs are interpreted in a qualitative way by looking at the graphs. However, both interrater and intrarater reliability for visually interpreting EEG-patterns are generally low (Benbadis et al. 2009, Grant et al. 2014). In my interviews, this was acknowledged as a challenge for cEEG-based prognostication:

The expectations are really high, while in daily practice... Don't forget, the average doctor isn't capable of interpreting this machine. [...] Our experience shows that even an average neurologist can hardly interpret. (Willems' intensivist)

In the ICUs studied, this issue was not that pressing; clinician-researchers (in all cases neurologists) specialized in cEEG were available. However, this definitely will not be the case when cEEG is implemented in non-research hospitals.

Additionally, it was clear from my observations that even experts go about interpretation in different ways. Dutch clinician-researchers would be called to the ICU to interpret the raw cEEG-patterns by the bedside, or look at them remotely in their offices. Either way, they visually assessed the activity at 12 and 24 hours after cardiac arrest. There is no standard procedure for this evaluation. The US clinician-researchers, in contrast, scroll through all data of the entire 12 hours that preceded their shift in an off site lab. Unusual patterns, spikes, and flat parts are studied more closely. Thanks to the addition of video monitoring, they can observe the corresponding activity at the bedside. Next, for all cases they take a 2 hour interval they consider representative and analyze it in detail. Finally, interpretations are discussed in detail as a team.

Clearly, the American procedure is very time-consuming, labour-intensive, and hard to implement in non-research hospitals. Dutch researchers are therefore seeking other solutions. They hope to reduce interrater and intrarater reliability through computer-support. Advantages of automated analysis are standardization and efficiency. Broadly speaking, there are two ways computers can be used to analyze the data from cEEG monitoring. One is as an aid to quantify differences among measurements, using pre-set rules, specified by human researchers (Tjepkema-Cloostermans et al. 2013). The other is based on machine learning techniques which allow the computer to identify regularities programmers did not anticipate. Thus, machine learning may potentially address the challenge regarding the comprehensiveness and precision of the list of patterns mentioned in the previous section. Recent publications show that machine learning techniques could further increase the ability to accurately predict neurological outcome after cardiac arrest, thus reducing uncertainty for an increasing number of patients (van Putten et al. 2017; Tjepkema-Cloostermans et al. 2017, 2019; Nagaraj et al. 2018). A major problem with deep learning techniques, however, relates to transparency (Mittelstadt et al. 2016, Silva & Kenney 2019). Human experts can check whether they come to the same conclusion in particular cases by looking at the patterns themselves but they would not know *how* the computer produced the prediction. Moreover, machine learning is prone to exacerbating self-fulfilling prophecies due to feedback loops. The concern of self-fulfilling prophecies³ requires in-depth analysis which I present in the final two chapters.

In summary, the work involved in classifying EEG-measurements requires expertise that will not be generally available. This poses serious challenges for the reliability of the resulting predictions, which researchers hope to tackle by developing computerized assistance. Depending on the technique used, such computer support may have challenges of its own.

³ Self-fulfilling prophecies are recognized to be a general problem in intensive care (Wilkinson 2009) and in neuroprognostication specifically, especially for patients after cardiac arrest (Geocadin et al. 2012, Gold et al. 2014).

5.1.3 Linking measurements to outcomes

The last step in producing a cEEG-based prognosis is linking EEG measurements and their development over time to neurological outcomes. Here, the first thing to note is that cEEG-based prognosis (like other prognostic technologies for this patient group) uses only two outcome categories: ‘good’ or ‘poor’. When researchers call surviving patients after 6 months or one year to check their wellbeing, they often register outcomes in terms of the 5 categories on the Cerebral Performance Categories-scale (CPC, see table 4.3 below). This scale classifies functioning on five levels, which are dichotomized for prognostic purposes. Recently, there have been pleas to shift to more fine-grained scales when measuring neurological functioning in research on outcomes after cardiac arrest (Balouris et al. 2015, Tong et al. 2016, Haywood et al. 2018). However, since it is not possible to link cEEG-patterns to the 5 categories of the CPC-scale, let alone to more fine-grained categorizations of outcome, results are reduced to two classes: good or poor outcome.

CPC	Activity Level	Outcome Classification
1 - Full recovery	Conscious. Can lead normal life and work. May have minor deficits.	Good
2 - Mild Cerebral Impairment	Conscious. Cerebral function adequate for part-time work in a sheltered environment, or independent activities of daily living.	Good
3 - Severe Cerebral Impairment	Conscious. Dependent on others for activities of daily living.	Poor (rougly since 2006); Good (before 2006)
4 - Coma or vegetative state	Not conscious. No interaction with the environment.	Poor
5 - Death	Brainstem dead or dead by traditional criteria.	Poor

Table 4.3 – Cerebral Performance Category Scale and outcome classification

Evaluating something in terms of ‘good’ or ‘poor’ clearly represents a value judgment, in this case pertaining to a patient’s future neurological state. Interestingly, as Sandroni and colleagues (2018, 151) show, and as discussed in previous chapters, the criteria internationally used to distinguish good from poor in neuroprognostic research have changed over time (see table 4.3). Whereas in the past CPC 1-3 were classified as ‘good’ outcome and CPC 4-5 as ‘poor’, CPC 3 in the scientific literature and clinical practice guidelines has gradually come to be perceived as ‘poor outcome’ (Sandroni & Nolan, 2015). As such, ‘poor outcome’ before 2006 was generally understood as referring to vegetative state and death;

'good outcome' usually referred to normal functioning and moderate and severe neurological disability. After 2006, most studies included severe disability in 'poor outcome'.

This particular change suggests a shift from 'regaining consciousness' as the main criterion distinguishing good and poor quality of life, to recovering sufficient cognitive and motoric capacity to allow independent daily living. This points to the question of who should determine what counts as 'good' or 'poor' quality of life, and how to draw the line between the two. Since the shift visible in research also fed into clinical guidelines, it represents at least a change in medical-professional norms. It is less clear to what extent this professional change reflects a broader shift in societal norms. The results of a national survey presented in the following chapter at least suggest that there may be a divide between clinician and public perspectives. Even if the shift reflected the norm, such a shift was probably more clearly present in some countries, rather than others. Surely, there is ample room for interesting historical research here.

In summary, even though cEEG-patterns enable reliable classification of more patients, the categories 'good' and 'poor' still cover a wide range in outcomes – as they do with existing prognostic technologies. Moreover, the boundary between the two is not self-evident and has changed over time. As discussed in the previous chapter, predicting how a patient will experience the neurophysiological outcome is far from straightforward, even with cEEG-results predicting what that outcome will be. Rather, the binary outcome measure obscures that outcomes vary in several respects, for example, with regard to mental and physical functioning. Moreover, it obscures that outcomes can be evaluated differently. While these remarks are not specific to cEEG, they do put the hopes raised by cEEG in perspective.

In conclusion,

- prognosis based on cEEG faces several limitations and uncertainties due to arbitrary pattern identification and variation in interpretations of patterns, which make its use far from straightforward
- in practice, these limitations and uncertainties are addressed by (1) developing extensive interpretation practices requiring expertise (and time) which is not widely available, or by (2) developing computer support, the workings of which may not be fully transparent to users and which may be hampered by feedback loops – all of which challenge the reliability of the technology
- cEEG, like other prognostic technologies in postanoxic coma, has an inherent normativity that is not transparent to users

5.2 Facilitating decision-making

I discussed how cEEG-monitoring produces predictions of 'poor' or 'good' quality of life. Importantly, results are often uncertain. My fieldwork suggests that, even

though cEEG definitely provides prognostic information for substantially more patients, the connection with decision-making for each prognostic result is far from straightforward. Below, I highlight why this is the case for each type of outcome respectively.

5.2.1 Responding to ‘poor’ outcomes

Decisions about whether or not to withdraw life-sustaining treatment requires a combination of normative criteria and predictive information. The former entails the question ‘What is a life worth living?’ The latter requires answering ‘What quality of life is to be expected for this particular patient?’. Above, I showed how cEEG-monitoring provides the second type of information but also implies a specific norm as to what is or is not a life worth living. To be clear, in no way does this imply that this specific norm is automatically applied to all patient cases. My observations clearly show that physicians make a dedicated effort to clarify how the patients themselves would value their life with particular limitations.

Legal declarations of which actions should be taken when the person in question can no longer express their wishes can be very helpful. However, since such advance directives are often missing, relatives are usually the most important source of information in this regard. The patient may have explicitly discussed his preferences. Mr. Reinhart, for example, had talked extensively about his views on a life worth living. His son recounted that any physical disability was a price his father would be happy to pay if it meant more time with his family. His father made it clear though, that if he would no longer recognize his wife or children, if he would lose the memories of who they were to him, they *had* to let him go. To him, that seemed a fate worse than death. His wife added that his wish made it impossible for her to keep him alive even if she wanted to. She felt that would be selfish. Such clear statements are often missing, however. Family and physicians then resort to reconstructing a patient’s preferences from their more general views of life and previous experiences.

It is clear that cEEG cannot, and does not aim to, solve the problem of how to infer which choices a patient would have made in this particular situation. However, regardless of how sure relatives are about what a patient would have wanted, connecting these views to a cEEG-based prognosis may be challenging. My observations indicate that both patients and relatives tend to think about (un)acceptable quality of life in different and in more nuanced terms than cEEG-monitoring allows for. This is clear in Reinhart’s case, who made a distinction between physical and cognitive or emotional capacities. Similarly, Willems insisted that interaction with others was most important to him. Losing the ability to express himself, to have meaningful exchanges, he no longer saw the point to go on living.

Both Reinhart and Willems thus had very specific damages in mind that to them would make continued living unacceptable. As seen in chapter 2, this nuanced way of thinking emerged quite frequently in the interviews and observations I conducted. Yet, even when referring to the underlying categories of the CPC and abstaining from the normative judgment of what is a poor and what is a good outcome, the CPC categories do not allow for distinguishing among different types of damage that are described by the patients above. CPC3, in particular, covers a very wide range of function losses. Thus, the type of outcome predictions offered by prognostic technologies, including cEEG, would not offer the qualitative information that my respondents are looking for.⁴ For now, the precision that respondents are interested in cannot be provided. In practice, this implies that medical professionals have to put a lot of effort in explaining what a poor prognosis actually predicts.

Summing up, when cEEG predicts poor outcome, this does not provide the type of information that allows linking the prognosis to patients' and relatives' views of a life worth living. The technology, like all other available prognostic tools, thus facilitates decisions about whether or not to withdraw life-sustaining treatment in a fairly limited way.

5.2.2 Responding to 'good' outcomes

One of the specific advantages of cEEG is that it offers early and robust prediction of *good* neurological recovery. The recently published new Dutch national guideline for postanoxic coma states that 'although this information does not yet yield treatment decisions, it is of great value to the treating physicians and the family of the patient' (Horn et al. 2019).

The value referred to concerns, first of all, the knowledge that the patient is likely to recover. This information justifies continuation of life-sustaining treatment, thus pre-empting the need to consider whether or not to withdraw it. This relieves both physicians and relatives from the heavy burden discussed in the previous section.

Second, my study showed yet another way to use good outcome predictions. Many medical respondents indicated they see it as their duty to provide sensible care. One respondent specified that aggressively treating infections or organ failures in a patient who may remain in vegetative state seems nonsensical to them. The clinical observations revealed that in case of a good prognosis, in contrast, intensivists find it self-evident to treat these ailments more aggressively.

⁴ I am not suggesting that these distinctions are necessarily informative for normative arguments about whether or not to discontinue someone's life support. My focus lies on the concerns and considerations that surface in practice, regardless of whether these lead to morally justifiable decisions.

In the same spirit, good news could result in more proactive care decisions aligning with the prediction, such as stopping the cooling and sedation and aim for early awakening. Several medical professionals suggested that research ought to be done on the benefits of early awakening of patients with a good prognosis.

In a nutshell, rather than 'just' being good news, cEEG's capacity to predict good neurological outcome has multiple values for decision-making. It justifies continuation of treatment and preempts the need to decide about the life and death of the patient. Moreover, the health professionals among my respondents indicate that there may be unexplored potential regarding treatment decisions for these patients.

5.2.3 Responding to uncertain outcomes

An additional, yet unexpected, effect of cEEG's capacity to predict good outcomes is that cEEG now also creates a more distinctive category of 'uncertain' measurements. Existing tests produce a prediction half of the time but only for patients with poor outcome. For the majority of cases they do not produce any prediction and for those patients, the advice is always to 'wait and see'⁵. While cEEG does not allow for a prediction in only half of *all* cases approximately, the meaning of the failure to provide a prediction is likely to be different in a number of ways.

First, the meaning of uncertain results in case of cEEG is qualitatively different because uncertain now implies 'neither clearly poor, nor clearly good', whereas before the lack of positive test result meant merely 'not clearly poor'. Even though physicians in these latter cases try to temper premature optimism by making explicit that lack of a positive test result is not necessarily a good sign, families nevertheless experience the message of 'wait and see' as positive. This is due to the impact a positive test result has, namely, withdrawal of life-sustaining treatment and, subsequently, the death of the patient. The fear and anxiety in anticipation of such bad news fall away, at least temporarily. Vanderbilt's wife, for instance, was told that if the SSEP-test predicted poor outcome, medical treatment had to be stopped. When that wasn't the case, she said: 'That was again a positive, very weakly positive, sure, but nevertheless, in the sense that the continuation of medical [treatment] could take place.' There is a signal, and even if it's hard to say what that means, there is reason for hope: at least it is not 'clearly poor'. With cEEG's ability to predict good outcomes as well, even if the advice for uncertain cases will still be to 'wait and see', it structures the anticipation of possible outcomes in a different way – for both physicians and relatives. The continuous character and the visibility of the EEG-measurements also contribute to this, as I will discuss in the next section.

⁵ Technically, this strategy can also be understood as a form of ICU-test (Vincent 2005) or time-limited trial of intensive care treatment (TLT) (Quill & Holloway 2011, Vink et al. 2018) but they were never communicated as such. Other studies confirm this is not unusual. Even when offered, TLTs are typically not sufficiently discussed (Schenker et al. 2013).

Second, the composition of outcomes in the ‘uncertain’ category changes depending on how many good and poor outcomes can correctly be identified. Although ultimate survival to discharge rates show substantial international as well as regional variation (Myat et al. 2018, Kundenchuk et al. 2015), chances of functional survival after successful reanimation and admission to the ICU, at least in the Netherlands, are more or less fifty-fifty (Tjepkema et al. 2015). This means that without prognostic information, outcome in all patients is uncertain. About half of these patients will have a poor outcome and the other half will have a good outcome. Once predictions can be made for *some* patients, the ratio of good to poor outcome expectancy for the remaining patients changes.

Before cEEG was implemented the SSEP-test was able to identify 30-48% of patients with poor outcomes (Zandbergen et al. 1998, 2006; Sandroni et al. 2018). Overall this implies outcome could be predicted for 15-24% of all patients. Let’s take the most optimistic view, which is that poor outcome can be predicted for nearly a quarter of the total population, leaving the prospects of the remaining three quarters of the population uncertain (see table 4.4). ‘Uncertain’ at that point means there is up to twice as much chance of recovery than there is a chance of poor outcome.

Prognostic ratio	Good/Poor outcome	Uncertain outcome (ratio good + poor)
No prognostic tools	0 % / 0 %	100 % = 50 % good + 50 % poor
SSEP	0 % / 24 %	76 % = 50 % good + 26 % poor
cEEG	25 % / 25 %	50 % = 25 % good + 25 % poor

Table 4.4 – Prediction ratio, outcome prediction (good/poor) versus no outcome prediction (ratio good/poor)

With the introduction of cEEG, however, filtering out both poor and good outcomes in equal measure means that chances of recovery for those with uncertain results are back to fifty-fifty. In essence, by filtering out both good and poor outcome, cEEG increases the ambiguity for patients in the uncertain category, returning it to the situation before testing. While these shifts in probabilities do not change the actual outcome of an individual patient, they can increase or decrease the clinical dilemma discussed in the previous chapter, making decision-makers more or less confident about what to do. In practice, all uncertain cases are considered as equally uncertain, and relatives are still told: ‘We don’t know, we must wait.’ However, as we’ll see in the next section, both predictive and uncertain measurements are prone to invite interpretation, depending on how visible the continuous measurements are. Their ultimate impact will depend on how and to whom uncertain results are made visible.

Finally, the question arises whether cEEG does not also filter out the very worst and the very best. Patients for whom cEEG does not provide predictive results ('not clearly good, nor clearly poor') *may* also be the patients with the most ambiguous *kind* of damage, for example, limited cognitive capacity, affecting their sense of personhood, while retaining motoric function. Since this is a hypothesis, further research into the long term outcome of patients for whom cEEG does not produce a clear prediction seems advisable. One cEEG researcher I spoke to confirmed they would be interested in conducting such research. If the hypothesis would be confirmed, this would justify the development of dedicated care approaches for this particular group of patients and their relatives.

Summing up, cEEG does not only predict good and poor outcome. It also creates a third group of patients, for whom the lack of cEEG results actually increases ambiguity and may decrease hope, even though treatment policy does not change. Pending further research, specific support for the relatives of patients in this group may be warranted.

In conclusion, with regard to facilitating decision-making, cEEG does both more and less than initially expected:

- Although it increases the reliability of predictions of poor outcome, it does not provide the detailed information stakeholders need for well-considered decision-making on the withdrawal of life-sustaining treatment.
- On the other hand, cEEG's predictions of good outcome *does* impact treatment and may enable even more directions for moving forward than currently acknowledged, for instance, by allowing early awakening or active neurostimulation.
- Finally, cEEG distinguishes a third category of patients for whom the anticipated future becomes more ambiguous than before the introduction of cEEG. As I will discuss in the next section, whether decision-making about the intermediate group requires dedicated support largely depends on how the information about this category is made available to relatives and physicians.

5.3. Facilitating communication

Aside from expecting cEEG to facilitate interpretation and decision-making, it seems reasonable to expect that communication with relatives about the patient and how to proceed will become easier once there is less uncertainty. Yet, fieldwork shows, once again, that actual impact on communication practices may be more complicated and ambiguous than first anticipated. The continuous character and bedside location of the EEG measurements in particular pose specific challenges. Two aspects deserve further discussion: the visualization of measuring results as well as the timing of communication.

5.3.1 Visualization of results

Any presentation of results will inevitably have an impact. As such, no presentation of information is ever completely neutral. Available information *will* be interpreted, and this influences decision-makers' expectations and, ultimately, their decision-making process. It is therefore crucial to consider how exactly specific presentations may help or hinder accurate interpretations and good patient care.

Traditionally, EEG measurements are visualized in graphs on a monitor. The use of continuous EEG implies that the monitor displays results in real time, like many other instruments at the ICU do. These may be the traditional graphs (considered 'raw data'), analyzed data, or both. The monitor can be at the bedside or elsewhere. The first thing to note here is that *if* the monitor is visible for anyone who happens to be at the ICU, as it was in one of the hospitals where I observed, this unavoidably invites interpretation. Willems' wife recounts:

I received no explanations. But I'm quite stubborn, so I go and look at the machine, to see, what does this do exactly? And I *got* that all we can do is wait at that point. They said so, too. But I also figured, if there's brain activity everywhere, even if just a little bit, then that is at least *something*. I mean, the moment I would see straight, flat lines, I would have rung the bell and told the doctor 'Hello, there are 3 straight lines here, that doesn't seem right to me. Either that thing is not connected very well, or there's no activity. But I would've appreciated getting a bit more explanation.

What interpretations are invited and what impacts they have depends on the way the monitor visualizes results. This part of the technology's design was still in development at the time I was doing fieldwork, so my colleagues and I reflected on the different options considered by the software developers and discussed them in a multi-stakeholder workshop. As mentioned above, tech developers aim to increase intra- and interrater reliability of EEG interpretation via quantitative analysis. This generates a number of choices regarding the representation of the results on the screen. They can be displayed graphically or with a number, in real-time or refreshed at pre-set moments. Interpretation can be facilitated by using colour coding. Key, here, is that each choice has specific implications.

Visualization furthermore inscribes who can use the information offered. The developers consider it desirable to extend the potential user group beyond the neurologist, as this facilitates implementation of the technology. Most hospitals have the equipment for EEG measurements, but they may lack sufficient neurological expertise. Simplifying the display of patterns means that other users besides the neurologist can make sense of the screen, like the intensivist, nurses, and other medical professionals - but also relatives.

In addition, simplification can be achieved in different ways, with different effects (see figure 4.1). When the display indicates that the pattern matches a pattern considered relevant (for example by colouring epileptiform patterns blue), it enables a first *descriptive* layer of interpretation by differentiating among a variety of patterns, without implying how to use this information. It still takes an expert to grasp what meaning these distinctions have and how to put them to use. When the visualization distinguishes patterns of a quantitative indicator (like the Cerebral Recovery Index: a number indicating chances of recovery (Tjepkema-Cloostermans et al. 2013, 2017)) indicating 'poor' and 'good outcome', it is *predictive* and already invites prognostication, which no longer requires specific expertise. Finally, designs using a traffic light model to distinguish outcomes tend to have a *prescriptive* effect. The colours convey 'stop', 'continue', or 'caution'. Now the technology *tells* the user what to do with the patient.

1. Example of descriptive visualization:

2. Example of predictive visualization:

3 & 4. Examples of prescriptive visualizations:

Figure 4.1: Comparison of visualization designs

The version I observed in practice combined display of raw, real time cEEG measurements with options for researchers to pull up other features like the graphical display of the development of the CRI. The developers were considering the use of red and green zones to distinguish good from poor results, with uncertain

results literally forming a grey area in between (see image 4 in figure 4.1). Such a design enables non-expert physicians to infer a prognosis, and thus facilitates broader use. A drawback is that it also invites interpretations by people not supposed to use it (not only relatives, but also nurses and some physicians).

When I discussed my analysis of design options with a panel of experiential experts, I received conflicting input. Two ICU-nurses criticized the idea of prescriptive visualization. They were worried that families seeing a curve in a red zone would jump to conclusions and warned that this would likely raise a lot of worries and additional questions, giving the nurses more work. In contrast, the neurologist and intensivist initially argued that the sooner the family is aware of the critical situation, the better. Yet, when it dawned on the intensivist that the *green* zone might create high expectations among relatives and potentially set them up for disappointment, the intensivist agreed that colored zones might not be the best way to go after all.

Another potential effect of visualization is that the 'uncertain' results become more visible than before and are more likely to be interpreted as having an 'intermediate' prognosis. This may reinforce relatives' (often already present) tendency to look for 'proof' of a patient's improvement⁶. In Vanderbilt's case, the family made a video of the patient. It showed the father lying in bed, reaching out to his daughter with his arm, seemingly attempting to stroke her face. The family showed this video to the intensivist, who then responded, seemingly convinced by the 'evidence', by saying: 'The truth is lying in bed'. Care for the patient was continued. The doctor's response returned repeatedly in the interviews with Vanderbilt's partner and children. It was the point they remembered as winning a negotiation process.

When the monitor visibly indicates an uncertain outcome, the explicit ambiguity invites a desire to postpone withdrawal of treatment and such negotiations may actually become more frequent. All medical professionals I spoke to agree that trust and rapport between the physician and family is crucial to reach agreement on treatment decisions. Yet, the visibility of uncertainty suggests that uncertain cEEG-based prognoses are a third explicit result next to the good and poor results rather than no result at all (merely implying 'yet to be determined'). As the example shows, explicit uncertainty can stall expert decision-making and lead to negotiations that undermine the relationship between physicians and family.

Visualization may also hide certain aspects from view. In particular, the source of the analysis is not visible on the screen. For instance, an interface using color-coding

⁶ In cases of prolonged disorders of consciousness, next of kin will often look for such proof even after explicit diagnosis of a poor outcome. The remaining uncertainty is then used to dismiss the validity of the test result altogether. (Shembs et al. 2019)

could represent a human steered analysis based on preselected features and predefined algorithms, or it could be based on deep learning. The interface would look exactly the same; the epistemic difference between the two is opaque to the user. The same goes for the CRI, which is represented by a number ranging between 0-1.⁷ Initially that number reflected as much as possible the experts' understanding (Tjepkema-Cloostermans et al. 2013). It was then further developed using machine learning training (Tjepkema-Cloostermans et al. 2017) and most recently is based on deep learning techniques (Nagaraj et al. 2018, Tjepkema-Cloostermans et al. 2019). Although these interpretative processes are significantly different, the CRI user is currently not made aware of this, as the visualization remains identical.

Clearly, different visualizations offer different functions, with different degrees of transparency for different people which impact how users understand, communicate and use the results. Since different stakeholders will value different functions of the technology, tradeoffs may be hard to avoid here. As a start, designers need to consider whether *everything* should be communicated to *everyone*, for *all three* outcome types, in the same way, at the same time.

As the family members I interviewed indicated repeatedly, the way bad news is delivered is key. Doctors and nurses in an unguarded moment commenting on a patient's state can cause traumatic experiences. However, medical professionals can be attentive and responsive to relatives' needs in a way a display never is. Likewise, when medical staff communicates positive news, they can contextualize and add nuance to the prediction, guarding recipients for disappointment, in ways that color-coded graphs on the monitor can't. Naturally, such responsiveness also depends on the communicator's ability to understand the prediction and the process that resulted in the prediction. As I have pointed out, some interpretative methods prevent this.

In conclusion, although there may be a strong desire to be transparent about prognostic information, the way this transparency is achieved is even more important. It is not just a matter of designing 'transparency' into the technology's user interface. Rather, the interplay of the design and implementation of cEEG on the one hand and communication with families on the other therefore needs careful consideration. Thinking about design and implementation in terms of shaping a testing architecture invites consideration of a broad set of elements that can be 'tweaked' to invite good professional behavior. It is important, however, to keep in mind the limitations of this way of thinking. Whether a medical professional is actually attentive and caring in their communication with relatives also depends on internal dispositions and capacities and the specific situation.

⁷ Deep learning does not currently allow a continuous calculation but only one at set times, currently at 6 hours and 12 hours after reanimation.

5.3.2. Timing of disclosure of prognostic outcome

At some point, the physician will discuss the patient's prognosis with the next of kin. One of the promises of cEEG monitoring is that it might provide a prognosis earlier than existing technologies. Research results thus far show that the technology may indeed provide a reliable prognosis within 12-24 hours after reanimation. The current Dutch clinical guidelines, however, advise to postpone communication of the prognosis till results from other tests are also available. This may be quite challenging if the monitor is visible for relatives and the interface easy to understand. Leaving this issue aside for now, cEEG raises the question whether it is desirable to inform relatives who have just heard that their loved one had a cardiac arrest about the prognosis as soon as possible?

Generally speaking, my respondents strongly emphasized the desirability of open communication, honesty, and transparency. Both family members and physicians expressed a preference for full disclosure of the prognosis as soon as information is available, even when a technology is still experimental, as Willems' wife insisted:

I certainly hope that the moment a doctor or researcher sees something, that they'll enter into conversation with the family or partner of the patient. 'Yes, this is experimental research, this isn't a given, of course. But we see some reason for worry.' That way, you at least create a bit of awareness, as in, keep in mind that someone doesn't get out of this the way you *think* they will.

Notice, however, that her reference to 'reason for worry' and preparing for unmet expectations imply that she is focused on transparency regarding *negative* information. Although early predictions of poor outcome are more controversial, since they may hasten the death of the patient, simply communicating the pessimistic prediction helps family members prepare for the worst and is therefore always appreciated. In my discussions with family and practitioners they seemed much less sure that communicating early predictions of good prognosis is desirable as well. Communicating an optimistic prediction, after all, doesn't have the same positive effect. Willems' wife elaborates:

Personally, I imagine it would be pleasant for the people around the patient to know 'we suspect some constraints or disabilities when he wakes up'. The suspicion of limitations means you'll be totally, super happy when the patient wakes up without limitations. If you have to wait until someone finally wakes up but then you're being told, 'there are limitations'... It's better to be prepared for bad news and get good news than to think, 'everything is going well' and then get smacked in the face later.

As long as a degree of uncertainty remains, family members indeed see an advantage to hearing an early negative prediction ahead of time which, if correct, prepares them for the pain and struggle that is to come, allows them some time for mental and practical preparation and to say goodbye to the patient. If the prediction turns out to be incorrect, that would simply be a source of happiness. However, such scenarios require delaying WLST based on the prediction, long enough for the patient to show signs of recovery. Family members further indicated that if that decision came too soon, they might reject the prognosis altogether.

Communicating an optimistic prediction early on, in contrast, raises expectations which can lead to disappointment if it proves incorrect, or if the patient dies from other causes than neurological damage. In these situations, the chance or time left to say goodbye may have been reduced significantly. I observed a similar response in my discussions with professionals. When asked whether they would be transparent about pessimistic preliminary findings early on, they only argued about the *best* way to deliver the bad news. When the preliminary findings were optimistic, however, they were reluctant to share their expectations, unless it could be made very clear that the prognosis was 'tentative'. This would guard the family against overly optimistic expectations about the outcome.

In contrast to what might be expected, then, early communication of a good prognosis is not necessarily appreciated. Different prognostic predictions seem to require different timing. To accommodate this requirement, both the design of the user interface and the communication policies that are part of the testing architecture need to be sufficiently specific. Like before, however, the personal disposition and capacities of the professional play a role that no testing architecture can fully 'inscribe'.

6. SUMMING UP AND LOOKING FORWARD

In view of the life or death decisions that are being made, reducing uncertainty regarding the future perspectives of a patient in postanoxic coma is important. Providing substantially more patients with a prognosis, as cEEG does, is therefore a desirable aim to pursue. The preceding analysis also shows, however, that the added value of cEEG in practice is less straightforward than one would expect. Fieldwork in two Dutch ICU's where cEEG was implemented for research purposes indicates that expectations of the new technology's ability to facilitate interpretation, decision-making, and communication, are not always met. Increased certainty does not necessarily produce 1) clear-cut information for medical professionals, 2) better guidance for decision-making, or 3) easier communication with relatives.

1. Regarding information provision, opening up the work and the choices underlying the production of cEEG results shows that cEEG faces several limitations and uncertainties regarding the identification and classification of patterns and how these are linked to outcomes. The different ways these issues of reliability are addressed, whether by additional resources or automation, pose their own challenges. Moreover, the technology has a built-in normativity that is rather crude and not immediately transparent.

2. This built-in normativity has implications for the way cEEG-results impact decision-making. Regarding poor outcome predictions, users feel the resulting prognosis still lacks the level of detail that they feel is required to decide whether or not to withdraw life-sustaining treatment. In contrast, good outcome predictions may actually have more action potential than has been acknowledged thus far. Most surprisingly, the technology also produces a new category of patients for whom the future becomes more, rather than less ambiguous.

3. With regard to communication, the early availability and potential accessibility of the prognostic information can foster, but also hinder, good communication with relatives. Although the design of cEEG may suggest that information is transparent to all, relevant nuances and assumptions are easily overlooked. Regarding timing, it is important to note that, even though cEEG provides a good prognosis earlier than a poor one, early communication of a poor prognosis may be more desirable than that of a good prognosis.

These observations are based on a limited set of observations and interviews in a limited number of settings in the Netherlands. I do not claim that use of cEEG will always work out in this way. On the contrary, the comparative investigation of the liminal use of cEEG in a US hospital shows that, in other settings, cEEG will be a different technology, with different impacts. Whether the different ways in which epistemic uncertainty is addressed are somehow related to differences in ethical culture needs further investigation, especially since the latter do not seem to impact the way normativity is inscribed. My findings do show, however, that this technology can bring along a number of social and ethical issues and challenges that need to be addressed to increase its acceptability. I hope that my analysis demonstrates how approaching design and implementation of cEEG as the building of a testing architecture helps to see how many different elements can be adjusted to avoid or address these challenges. It also helps to become aware of their interdependency. At the same time, it is important to remain aware that adjusting this architecture is not a panacea to prevent inattentive and uncaring professional behavior.

When further developing the technology's overall architecture, attention should be paid to safeguarding reliable interpretations: how to align the design of the technology, the working procedures to interpret measurements, and arrangements as to who is allowed to interpret. This includes careful consideration of the (dis)advantages of human versus machine-supported versus machine-based interpretation. Different decisions will come with varying trade-offs for transparency.

Importantly, the user interface must be designed and implemented in such a way that it allows for human-mediated ways and multiple timings of communicating the prognosis with relatives. For some practices, this may mean that relatives are prevented from seeing the screen altogether.

In addition, efforts must be made to nuance outcome measures, or (as long as this is not possible) transparency about what outcome categories imply must at least be increased, to allow for open discussion and for more informed decision-making. Investigating whether cEEG predictions are filtering out the very best and very worst outcomes by examining the uncertain category is thought to be a worthwhile endeavor in that regard.

Finally, furthering research on what additional proactive care can be offered to patients with a good prognosis, as well as what tailored support can be offered to patients with an uncertain (or 'grey') prognosis and their relatives, would increase the benefits of cEEG-based prognosis.

Chapter 5

Preferences for prognostication of coma after cardiac arrest:

the Dutch societal perspective on physiological and normative futility

ABSTRACT

Early outcome prediction in intensive care is an increasingly common practice for patients in coma after cardiac arrest. In the Netherlands, early neuroprognostication impacts futility judgments that inform decisions on (dis)continuation of life-sustaining treatment. The objectives of the study presented in this chapter were to elicit public preferences for prognostic information, decisions to withdraw life-sustaining treatment (WLST), and perceptions on acceptable quality of life after coma. A representative sample (n=500) of the adult Dutch population participated in a web-based survey with closed and open-ended questions and three discrete choice experiments. The results of this study indicate that current decisions based on predictions of physiological and normative futility are not fully representative of Dutch preferences. Regarding physiological judgments of futility, half of the study population considers WLST in patients in coma after cardiac arrest based on uncertain physiologic prognostic information unacceptable. With regard to normative judgments of futility, more than half of our study population prefers to stay alive in states which are considered a poor outcome in medical literature, research and practice. This indicates a need for a closer look at the practice of WLST based on prognostic information in the Netherlands, to ensure responsible use of novel prognostic tests.

The Dutch survey presented in this chapter is included in the comparative article currently under review as **van Til J, Bouwers-Beens E, Mertens M, Boenink M, Groothuis-Oudshoorn C, Hofmeijer J (2021) Prognostication of patients in coma after cardiac arrest: public perspectives**. This chapter originates from the same authors with Mertens contributing to the design of the study, the interpretation of the data, and the revising of the manuscript.

Preferences for prognostication of coma after cardiac arrest:

The Dutch societal perspective on physiological and normative futility

1. INTRODUCTION

The incidence of out-of-hospital cardiac arrest (OHCA) in the Netherlands is about 1 in 1000 inhabitants (de Vreede-Swagemakers et al. 1999). A meta-analysis by Yan et al. (2020) concluded that the global average rate of survival to hospital admission is about 22%, and the rate of survival to hospital discharge is about 9%. This means that, of the patients that arrive at the hospital alive, little over half will survive the first 3 months. Moreover, a systematic review by Moulaert et al. (2009) concluded that about half of survivors suffer from some type of cognitive impairment. An observational study performed in the Netherlands showed that between 15% and 30% of surviving patients had moderate to severe cerebral disability (Blom et al. 2014).

Outcome prediction of patients in coma after cardiac arrest is included in national and international guidelines (Sandroni 2014). Along with clinical examination, prognostic tests may guide decisions with regard to further treatment. In patients in whom a poor outcome is expected, life-sustaining treatment can be deemed medically futile and hence withdrawn. Studies have shown that in the majority of patients in coma that die after cardiac arrest, death occurs after the decision to withdraw life-sustaining treatment (Geocadin et al. 2006, Mulder et al. 2014, Grossestreuer et al. 2017). Therefore, it is extremely important to correctly identify patients with good and poor outcomes.

In outcome prediction, multimodal testing is recommended (Rossetti et al. 2016). Available tests include the somatosensory evoked potential (SSEP) and the pupillary and corneal light reflex test (Rossetti et al. 2016, Zandbergen et al. 2000). These tests can be performed from 48-72 hours after cardiac arrest, and help identify up to 48% of patients with a poor outcome (Zandbergen et al. 2006, Sandroni et al. 2014). Consequently, despite these tests, the prognosis of a vast majority of patients in coma remains uncertain in the first days to weeks after cardiac arrest. Previous studies have found that uncertainty about prognosis is perceived as burdensome by and to family members of patients in coma after cardiac arrest (Smith et al. 2013). Recently, predictive values of continuous electroencephalography (cEEG) have been established (Hofmeijer & van Putten 2016). In contrast to other tests, cEEG can predict both good and poor outcome (Admiraal et al. 2020). As a consequence,

cEEG can provide prognostic information within 24 hours after cardiac arrest in about half of the patient population (Sondag et al. 2017, Ruijter et al. 2018, Horn et al. 2019). As of mid-2019, cEEG is included in the Dutch Guidelines for outcome prediction of coma after cardiac arrest (Horn et al. 2019) and in the European Guidelines as of 2021 (Nolan et al. 2021).

Even though it is expected that outcome predictions of patients in coma after cardiac arrest are reliable with multimodal testing, the risk of an inaccurate prediction is never zero (Rossetti et al. 2016). Furthermore, in addition to remaining physiological uncertainty about the neurological outcome and thus the physiological futility of treatment, there can be normative uncertainty about which physiological outcomes are worth pursuing. When prediction of a poor outcome is followed by WLST, it is impossible to verify the prediction of both physiological and normative futility judgments as a result of the self-fulfilling prophecy: if treatment is withdrawn because of a poor prognosis, the patient will almost always die (Wilkinson 2009, Geocadin et al. 2012, Mertens et al. 2021).

In light of these concerns and the new ability to perform earlier outcome prediction of patients in coma after cardiac arrest, with the consequent far-reaching decisions on continuation or withdrawal of life-sustaining treatment based on test outcomes, questions about timing and accuracy of outcome predictions have arisen. The objectives of the study were to elicit public preferences for outcome prediction, the decision to withdraw life-sustaining treatment, and perceptions on acceptable long-term quality of life for patients in coma after cardiac arrest, including trade-offs between likelihood, timing, and accuracy of prognostic test information. As such, these public preferences can bring to light the existing views on physiological and normative uncertainty and help assess the current physiological and normative futility judgments used in research and practice.

2. MATERIALS AND METHODS

2.1. Study Design

This study was a structured web-based survey, which consisted of closed-ended questions and three discrete choice experiments. The study protocol for this study was approved by the Ethical Review Board of the University of Twente and the study was exempted from medical ethics review. Verbal informed consent to participate was obtained prior to the interview and audio recorded. The study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of the University Twente.

2.2 Survey instrument

Background information - Before starting the actual survey, respondents were asked for their age (in years) and gender (male/female) to include a representative sample of the Dutch population-based on these characteristics. Further background information, which was collected at the end of the survey, included educational level (low/middle/high), religion (Christian/Muslim/other/non-religious), migration background (yes/no), and experience with having a close relative or friend with coma after cardiac arrest (yes/no).

Perspectives on quality of life - In medical literature, outcome of coma after cardiac arrest is usually classified according to the cerebral performance category (CPC), a five-point scale ranging from 1 to 5, with 1 indicating full recovery and 5 death. CPC-3 has been the grey area with regard to what is considered a poor prognosis (Sandroni et al. 2018). In this study, three health state descriptions within the range of CPC-3 outcomes and one for CPC-4 were formulated. Descriptions of physical and cognitive abilities were based on the Cerebral Performance Categories – Extended (CPC-E) Scale (Balouris et al. 2015). The health states were developed in collaboration with clinical and ethical experts. Validation of the health states took place in a pilot study. After being presented with each of the health states, respondents were asked to indicate whether they would want to stay alive if the outcome would be as described on a 4-point Likert scale (ranging from definitely yes to definitely no).

Preference for prognostic information – First, respondents were asked to imagine having a close family member in coma after cardiac arrest. Preferences for receiving information on the predicted outcome in case of good and poor prognosis, and preferred timing of information on poor prognosis, were elicited by directly asking respondents whether and when they would like to receive this information. Second, discrete choice experiments (DCEs) were used to elicit the trade-offs respondents make in their desire to receive prognostic information. A DCE is a health preference method, which is especially suited to elicit respondent trade-offs between attributes of health innovations that are thought to determine preference (Ryan 2004). Different possible outcomes are formulated for each attribute, which are called levels. In a DCE the respondents are repeatedly asked to choose between two hypothetical scenarios, which consist of different levels for the different attributes of the innovation.

In this study, three different DCEs were developed according to the good research practice guidelines (Bridges et al. 2011). In DCE1, respondents were asked to trade between timing and accuracy of prognostic information predicting good outcome. In DCE2, respondents were asked to trade between timing and accuracy of prognostic information predicting poor outcome. Accuracy estimates were based

on literature (Rossetti et al. 2016). In both DCEs, respondents were also allowed to opt out of receiving information. The order of DCE1 and DCE2 was randomized over respondents.

In DCE3, respondents were asked to trade between the likelihood that prognostic information can be given and the accuracy of prognostic information predicting good and poor outcome. Accuracy and likelihood estimates were based on the variation between existing tests (Rossetti et al. 2016). In this DCE, no opt-out was offered. In Table 5.1, the attributes and levels per DCE are presented. Before starting with the DCE, the attributes and levels were explained to participants and written instructions on how to complete a DCE choice-task were given, along with an example of each task.

For DCE1 and DCE2, the full factorial design was used (Reed et al. 2013). For the 3 attribute 4 level DCE (DCE3), a fractional factorial design was developed with 48 unique choice-tasks. Choice tasks with dominant scenarios were excluded. In all DCEs, level balance and orthogonality were maintained. All DCEs were blocked over 12 versions of the survey. Each respondent received four choice-tasks of DCE1, four choice-tasks of DCE2, and four choice-tasks for DCE3.

Attitude toward withdrawal of life-sustaining treatment (WLST) - Direct questioning was used to elicit attitudes towards WLST. First, respondents were informed about the current clinical practice of WLST. Then, respondents were asked to indicate whether they agree with the decision to withdraw life-sustaining treatment in case of poor outcome (yes/no/depends); whether they would accept a decision to withdraw life-sustaining treatment in case of uncertainty about prognostic information (yes/no); how much uncertainty would be acceptable in a decision to withdraw life-sustaining treatment (predefined categories ranging from 0.1-1.0%); and what would be the best timing for WLST after the decision is made (predefined categories ranging from 1-7 days after prognosis/other).

Involvement in decision making - Preferences regarding involvement in decision making were elicited using direct questions. Respondents were asked who should initiate the conversation on WLST; and who should be responsible for making the decision regarding WLST.

The survey was pilot-tested in two phases among a convenience sample of respondents, which were recruited through snowball sampling. In the first phase, ten 'think aloud' tests were performed to verify comprehension, readability, and feasibility of the survey. In this phase, the researcher was present and respondents were able to give feedback. In the second phase, 56 respondents completed the revised survey online to ensure validity and feasibility.

Attribute	Level 1	Level 2	Level 3	Level 4
DCE1 Good Prognosis				
Timing of test result	12h after OHCA	24h after OHCA	48h after OHCA	72h after OHCA
Accuracy of good outcome prediction	80 %	90 %	95 %	98 %
DCE2 Poor Prognosis				
Timing of test result	12h after OHCA	24h after OHCA	48h after OHCA	72h after OHCA
Accuracy of poor outcome prediction	80 %	90 %	95 %	98 %
DCE3 Availability and accuracy of prognostic information				
Likelihood of prognostic information	20 %	30 %	50 %	60 %
Accuracy of good outcome prediction	80 %	90 %	95 %	98 %
Accuracy of poor outcome prediction	80 %	90 %	95 %	98 %

Table 5.1 - Attribute and levels

2.3 Study population

The study was conducted among a sample of the Dutch general population, which was recruited through a market research agency, in September and October 2019. Respondents were eligible if they were at least 18 years old, living in the Netherlands, and were able to provide informed consent. Based on the Orme's Rule of Thumb, the minimum sample size required for the DCE was 250 (de Bekker-Grob et al. 2015). We oversampled to 500 respondents to study the effect of background characteristics on preferences. To reach a sample representative for the Dutch population, quotas were maintained for gender and age in the recruitment of respondents.

2.4 Statistical analysis

Descriptive statistics (frequencies) were used to examine background characteristics and responses to direct questions and rating scales. In order to assess the representativity of the study sample for the Dutch population, differences in age, gender, educational level, religion, and migration background between the study population and the Dutch population were tested using chi-square tests. A non-parametric Friedman test and a post-hoc Wilcoxon test were used to study the impact of background characteristics on the perception of acceptable quality of life.

Conditional logistic regression analysis was used to analyze the data from the DCEs. First, all attributes were treated as categorical variables. A clear linear preference ordering was established for the attributes accuracy of prognostic information predicting poor and good outcome, and the likelihood of any prognostic information, so these were then included as continuous variables. Timing was included as a categorical variable. Attribute weights were calculated using profile-based normalization as described by Gonzalez (2019). The odds ratio was calculated by taking the exponent of the β -coefficient from the regression analysis.

Ordinal regression analysis was used to test association between age, gender, educational level, religion, relationship status, parental status, experience with patients in coma after cardiac arrest, or own perceived health and the acceptability of a decision to withdraw life-sustaining treatment and the perception of acceptable quality of life. All analyses were performed using R (version 3.6.0, www.r-project.org) or Excel. The significance level was set at 0.05.

3. RESULTS

3.1 Background Characteristics

In total, 500 respondents completed the survey. The mean time to complete the survey was 15 minutes. The background characteristics of the respondent sample can be found below, in Table 5.2.

Characteristic	Level	n=500
Age		49 (18-88; 18)
Gender	Male	258 (52 %)
	Female	242 (48 %)
Educational level	Low	112 (22 %)
	Medium	190 (38 %)
	High	198 (40 %)
Migration background	No	425 (85 %)
	Yes	75 (15 %)
Relationship	Yes	347 (69 %)
	No	153 (31 %)
Children	Yes	302 (60 %)
	No	198 (40 %)
Religion	No religion	282 (56 %)
	Christianity	193 (39 %)
	Islam	9 (2 %)
	Other	16 (3 %)
Experience with patients in coma after cardiac arrest	Yes	121 (24 %)
	No	379 (76 %)
Perceived health		7.4 (1-10; 1.50)

Table 5.2 - Sociodemographic and clinical characteristics of respondents (n=500). Data are reported as n (%) or mean (min-max; SD).

3.2 Perspectives on quality of life

Fewer respondents prefer to stay alive with increasing severity of the health state description (see Table 5.3). However, thirty-five respondents (7%) are certain, and 40 (8%) are fairly certain they want to stay alive in vegetative state. A total of 25% of respondents are certain or fairly certain they want to stay alive in a state in which they are conscious, but with severe physical and cognitive impairments (CPC3). Fewer people indicate they (probably) want to stay alive in a conscious state with only physical (56%) than with only cognitive impairments (61%).

Younger respondents are more likely to indicate they want to stay alive in a vegetative state (OR 1.04, $p < 0.001$) or in a conscious state with both severe physical and cognitive impairments (OR 1.03, $p < 0.001$) than older respondents. A 20-year difference between the age of respondents makes it about twice as likely that the younger person wants to stay alive in a vegetative state.

	Want to stay alive	Probably want to stay alive	Probably do not want to stay alive	Do not want to stay alive
Vegetative state	35 (7 %)	40 (8 %)	141 (28 %)	284 (57 %)
Conscious with both severe physical and cognitive impairments	40 (8 %)	84 (17 %)	184 (37 %)	192 (38 %)
Conscious with severe cognitive impairments	110 (22 %)	196 (39 %)	128 (26 %)	66 (13 %)
Conscious with severe physical impairments	94 (19 %)	184 (37 %)	138 (28 %)	84 (17 %)

Table 5.3 - Willingness to live for predicted long-term outcomes of coma after cardiac arrest (n=500). Data are reported as n (%).

Women are about four times more likely to indicate they don't want to live in a vegetative state (OR 3.618, $p < 0.001$) and two times as likely to indicate they don't want to live in a conscious state with both severe physical and cognitive impairments (OR 1.80, $p = 0.01$) or in a conscious state with severe cognitive impairments (OR 1,822, $p = 0.00$) than men.

Christians are about half as likely to indicate they don't want to live in a vegetative state (OR 0.457, $p = 0.01$), in a conscious state with both severe physical and cognitive impairments (OR 0.439, $p = 0.00$), and in a conscious state with only severe physical impairments (OR 0.407, $p < 0.001$) than non-religious people. Also, Muslims are about ten times less likely to indicate they don't want to live in

a conscious state with both severe physical and cognitive impairments than non-religious people (OR 0.104 $p = 0.04$).

Respondents with a better self-reported health status are more likely to want to stay alive in a vegetative state (OR 0.796, $p = 0.03$), while respondents without experience with patients in coma after cardiac arrest (OR 1.960, $p = 0.02$) are about twice as likely to prefer death to living in a vegetative state than those with experience.

Finally, respondents with a high educational level are about twice as likely to prefer death to being alive in a conscious state with both severe physical and cognitive impairments (OR 1.893, $p = 0.04$).

What is perceived as an acceptable quality of life is not significantly associated with relationship status or whether or not respondents have children for any of the health states.

3.3 Preference for prognostic information

In this part of the survey, good outcome was defined as good recovery or moderate disability (CPC 1 or 2). Poor outcome was defined as either death, vegetative state, or severe disability (CPC 3-5). This dichotomization reflects current research and care practice (Sandroni et al. 2018).

The results of this study indicate that the majority of the respondents would like to receive prognostic information both if it predicts a good outcome (92%) and when it predicts a poor outcome (89%).

When respondents were asked to choose between two scenarios that offered prognostic information predicting good outcome, 52 respondents (10.4%) opted out of receiving the information at least once (DCE1). Fourteen respondents (2.8%) always opted out of receiving information. The most frequent reason for opting out was the (in)accuracy of the prognostic information.

Eighty-seven respondents (17.4%) opted out of receiving the poor outcome prediction at least once, while 25 respondents (5%) always opted-out out. Again, the reason that was most frequently mentioned was the (in)accuracy of the information.

Likewise, preferences with regard to timing of information that predicts poor outcome varied between respondents. Of the respondents, 29% prefers to receive prognostic information predicting poor outcome at 24 hours, 33% at 48 hours, and 30% at 72 hours after cardiac arrest.

In case of prognostic information predicting good outcome, respondents attached most importance to the improvement of the accuracy of the test (91%). With regard to timing of information, respondents preferred prognostic information predicting good outcome at 48 hours instead of 12 hours after admission ($p=0.02$). In case of poor prognosis, respondents only attached importance to the improvement of the accuracy of the test (93%). There was no significant impact of timing on preference for receiving a test result (see Table 5.4).

Attribute and levels	Coefficients	Odds Ratio	Standard Error	p-value	Attribute Importance
Good prognosis					
Timing of test results*					9 %
12 hours after admission	0	1.00	-	-	
24 hours after admission	0.186	1.20	0.098	0.06	
48 hours after admission	0.257	1.29	0.112	0.02	
72 hours after admission	0.074	1.08	0.096	0.44	
Accuracy of test result (80-98 %)	0.144		0.007	0.00	91 %
> 10 %		4.22			
Poor prognosis					
Timing of test results*					7 %
12 hours after admission	0	1.00	-	-	
24 hours after admission	0.108	1.11	0.088	0.22	
48 hours after admission	0.101	1.11	0.101	0.32	
72 hours after admission	0.108	1.11	0.087	0.22	
Accuracy of test result (80-98 %)	0.081		0.005	0.00	93 %
> 10 %		2.25			
Availability of receiving prognosis (20-60 %)	0.028		0.002	0.00	41 %
> 10 %		1.32			
Accuracy of good prognosis (80-98 %)	0.050		0.005	0.00	33 %
> 10 %		1.65			
Accuracy of poor prognosis (80-98 %)	0.040		0.005	0.00	26 %
> 10 %		1.49			

Table 5.4 - Coefficients following from regression analysis and the relative importance of attributes
*categorical variable

Respondents attached most importance to the increased likelihood of receiving any prognostic information from 20 to 60% (41%) compared to the other test characteristics. An increase in accuracy of prognostic information predicting good outcome from 80 to 98% (33%) was considered slightly more important than the same increase in accuracy of prognostic information predicting poor outcome (26%).

The odds ratios indicate that the likelihood that a prognostic test is preferred is most influenced by a 10% higher accuracy of positive prognostic information (OR 1.65), and least by a 10% increase in likelihood of receiving any prognostic information (OR 1.32).

3.4 Attitude toward withdrawal of life-sustaining treatment

According to 65% of the respondents, WLST is allowed in case of a prediction of poor outcome. Twenty-one percent (21%) feels WLST should not be allowed and 14% indicates that whether or not WLST is allowed depends on the accuracy of the test or the patient's wishes. When informed about the fact that prognosis is inherently uncertain, WLST is considered acceptable by only 50% of the respondents. Among the respondents that do accept WLST in case of uncertainty, the median accepted uncertainty is 0.5% [range 0.1-1%] indicating inaccurate prognosis in 1 in 200 patients.

Two-hundred-and-seventy-five respondents (55%) indicate WLST is allowed within the first three days after receiving prognosis. One hundred and twenty-five respondents (25%) indicate that WLST is allowed between the fourth and seventh day. The remaining 20% preferred to wait longer, or did not allow WLST.

The odds of a 30-year-old accepting WLST are about half of that of a 65-year-old respondent (OR 0.5, $\beta = -0.020$, $p = 0.01$). In contrast, respondents who have experience with coma after cardiac arrest are about twice as likely to accept WLST in case of poor prognosis (OR 2.1, $\beta = 0.73$, $p = 0.01$) compared to respondents without experience, also when prognostic information is not 100% accurate (OR 2.2, $\beta = 0.77$, $p < 0.001$).

When asked about important considerations regarding WLST in case of poor prognosis, most frequently mentioned from the prespecified list were the wishes of the patient (73%), expected quality of life of the patient (72%), and accuracy of the prognosis (59%). Costs of treatment were considered relevant by only 11% of the respondents.

3.5 Involvement in decision making

Sixty-two percent (62%) of the respondents prefers the clinician to initiate the conversation about possible WLST. Forty-eight percent (48%) of the respondents prefers the decision to WLST to be a shared decision between the medical team and the family, while 23% indicated that the medical team should be responsible and 29% indicated that the family should be responsible for making this decision. When there is disagreement in joined decision-making about half of the respondents indicated the family should have the final say.

4. DISCUSSION

The objective of the current study was to elicit public preferences for outcome prediction, the decision to withdraw life-sustaining treatment, and perceptions on acceptable long-term quality of life for patients in coma after cardiac arrest.

The study results demonstrate that the Dutch public considers a higher accuracy of prognostic information more important than earlier availability. This finding is in line with the results of a previous study by Steinberg et al., which found that health care providers from different countries favor more accurate prognostic tests to faster test results (2019). The results of our current study show that increasing the odds of prognostic information is valued by the respondents, but that test accuracy is a stronger predictor of perceived test value than the likelihood that any prognostic information (both positive and negative) is available.

Although WLST based on a predicted poor outcome is common practice in the Netherlands, about one-third of our study population feels WLST should not be allowed when a poor outcome is predicted and half of the population does not accept WLST with some degree of remaining prognostic uncertainty. To Dutch health care professionals working in this field, this finding is surprising [personal communication]. In their experience, most families support the decision to withdraw life-sustaining treatment when a poor outcome is expected. The results of our study also indicate that experience with coma after cardiac arrest does influence preferences. However, the findings are somewhat contradictory. Respondents with experience with patients in coma after cardiac arrest are more likely to accept WLST in case of a poor prognosis, but when reporting their personal perspective on different quality of life states, they are also more likely to prefer to continue living in a vegetative state. In the interpretation of these outcomes, we must take into account that actual numbers of respondents with experience were low and that written descriptions were used to elicit perspectives on quality of life and attitude towards WLST.

Over the years, perceptions on acceptable quality of life after coma have changed, and a more narrow view on acceptable quality of life was adopted in outcome evaluation after 2006 (Nolan et al. 2015). In current literature, poor outcome after cardiac arrest is defined as a score of three to five on the Cerebral Performance Categories (CPC) (Sandroni et al. 2018). The results of this study show that there is considerable variability in the extent to which respondents prefer death to being alive in a health state that would be considered poor according to this definition. About half of the study population perceives a conscious state with severe physical or cognitive impairments (CPC 3) as a health state in which they would prefer to live. The results of our study align with those of a study among health professionals by Friberg et al. (2015), and suggest that using only the CPC categories 4 and 5 as poor outcomes would be more in line with Dutch public perspectives on acceptable quality of life.

5. STRENGTHS AND LIMITATIONS.

A strength of this study is that a relatively large sample of the Dutch population was asked for their attitude and preferences towards prognosis and WLST for patients in coma after cardiac arrest. The sample included in this study is representative of the Dutch population in terms of age, gender, educational level and religion. There is a slight underrepresentation of respondents with a migration background, which might be due to the recruitment through online panels, the Dutch language of the survey, or both.

While the results of this study offer interesting insights into the public perspective on management of patients in coma after cardiac arrest, we also have to consider its limitations. We used a descriptive format to describe the long-term outcomes. The health state descriptions were based on the CPC and the CPC-extended. Previous studies criticized the CPC for having poorly defined criteria, the fact that some criteria include multiple domains, and that the scale has poor connections to measures of quality and disability of life (Hsu et al. 1996, Raina et al. 2008). By definition, written descriptions are always a reduction of reality.

However, the more important question is whether this reduction is directional, or favors a certain interpretation. In the design phase of our study, multiple clinical and ethical experts were involved in the design of the scenarios. We carefully considered presenting balanced words and excluding any words which suggested value judgments. We witnessed opposing reactions to the scenarios from health care professionals in the field. Some clinical experts saw them as too positive, while others found them too steering towards the negative. Therefore, while it is obvious that the descriptions of health states as we used them leave room for interpretation, we have no reason to think that we steered answers in a specific direction.

Similarly, any descriptive format of a health state will be interpreted differently by each survey respondent (or clinician), just like seeing a close relative in coma will be a different experience for each observing family member. While we feel that the results of our study do adequately reflect the diverging opinions of the Dutch public about prognosis of patients in coma after cardiac arrest, we cannot conclude from our current study whether the public preferences derived from our theoretical scenarios can be extrapolated towards preferences of families of actual patients in coma admitted to intensive care units. We do think that further research into the experiences and beliefs of family members of actual patients in coma on intensive care units, family members of patients that have died after WLST, and surviving patients would add valuable information to the discussion on outcome prediction of, and WLST in patients in coma after cardiac arrest.

6. CONCLUSION

The results of this study indicate that the highest valued attribute of a new prognostic test in patients in coma after cardiac arrest is its ability to offer prognosis in patients for which this information is currently not available. Study participants are willing to accept lower accuracy of prognostic information if their likelihood of receiving *any* prognostic information increases, but they are not willing to trade accuracy of the test results to receive information sooner.

However good the prognostic tests will become, some degree of uncertainty will always be present. The results of this study indicate that, regarding predictions of physiological futility, half of our study population considers WLST in patients in coma after cardiac arrest based on uncertain prognostic information unacceptable. Regarding predictions of normative futility, more than half of our study population prefers to stay alive in states which are currently considered a poor outcome in medical literature, research, and practice.

The results of this study indicate a need to have a closer look at the practice of WLST based on prognostic information in the Netherlands, to ensure responsible use of novel prognostic tests. Further study could focus on the question to which extent the public preferences derived from this study are in line with preferences of family members of actual patients in coma after cardiac arrest admitted to intensive care units, and on how communication between doctors and family members on the ICU influences acceptance of WLST.

Chapter 6

When hidden mistakes impede learning...

taking practical and epistemic responsibility for self-fulfilling prophecies

ABSTRACT

The potential for self-fulfilling prophecies (SFPs) in medicine is well acknowledged. From placebo and nocebo effects in medication and diagnosis to the self-fulfilling prophecy in prognosis and intensive care, scholars have pointed out the self-fulfilling impact that assertions about the effect of treatment or condition of the patient can have. Understandably, the interest focuses on predictions that change the effect of treatment or outcome of the patient. Unfortunately, the focus on reflexive predictions that have an altering effect has led practitioners as well as scholars to misunderstand both the mechanism of SFPs and the normative and epistemic issues that SFPs really entail. In this chapter, I describe how SFPs in medicine work and go on to challenge the current concerns by emphasizing the problems that SFPs pose regardless of whether or not they change the outcome. First, because SFPs are always true, the quality of the prediction is likely never called into question. However, because we cannot distinguish between *transformative SFPs* that change the outcome and *operative SFPs* that do not, one cannot counterfactually check to what extent the basis of the prediction was correct, even if one tried. In other words, because they entail their own fulfillment, SFPs do not produce reliable feedback. Not being able to identify whether or not the basis of the prediction was accurate or flawed, means one cannot learn from failure and one should not learn from 'perceived' success. This is the central epistemic failing of SFPs. Second, rather than SFPs necessarily allowing agents to evade responsibility for the practical outcomes they create—which may be inevitable, the bigger problem with SFPs regarding responsibility is that they allow agents to evade responsibility for their failure to learn.

This chapter partially adapts to a medical context material from the general account on SFPs developed in collaboration with Owen C. King. The general account is currently under review as **King OC & Mertens M (2021) Self-fulfilling Prophecy in Practical and Automated Prediction.**

When hidden mistakes impede learning...

taking practical and epistemic responsibility for self-fulfilling prophecies

1. INTRODUCTION

Self-fulfilling prophecies (SFPs) have long fascinated scholars and practitioners in many fields, not just medicine. In medicine, the placebo effect is one of the most commonly acknowledged SFPs and it has become the standard quality assurance for trials in medical research. While the placebo's effect is positive, improving the patient's health, pessimistic self-fulfilling prophecies have a negative effect. The latter are particularly worrisome in prognostic and early-diagnostic practices as the pessimistic predictions can be the cause of the eventual poor outcome. Whether their effects are positive or negative, however, SFPs have received the kind of attention they have, precisely because of their ability to change things for better or worse. If the mere making of a prediction can make it so, when *not* making the prediction would *not* make it so, how can one be anything but fascinated?

Unfortunately, the focus on the self-fulfilling prophecies that have an altering effect has led practitioners as well as scholars to misunderstand both the mechanism of SFPs and the normative and epistemic issues that SFPs really entail.

To remediate these limitations, this chapter presents an empirically informed theoretical analysis of self-fulfilling prophecies, especially in medical prognosis, that takes into account *all* SFPs, also those that do *not* change the outcome of the prediction. SFPs that alter the outcome, we call *transformative*. Those that do not, we call *operative*. The inclusion of operative SFPs is crucial in understanding the mechanisms behind self-fulfilling prophecies. Only when one fully understands how SFPs work, can one understand what is actually problematic about both optimistic and pessimistic predictions that self-fulfill.

While the analysis presented here applies to many kinds of SFPs in different fields (including examples of SFPs in education and predictive policing)¹, I will, in this chapter, lay out the different components focusing solely on the case of medical prognosis, and most often, on a case of neuroprognostication. The basic scenario is the following:

Consider Chris, an unconscious coma patient in intensive care. Suppose that, according to tests of Chris's brain activity, he is predicted to have a 'poor outcome',

¹ I discuss SFPs in practical and automated prediction more generally elsewhere (King OC & Mertens M, currently under review).

which could be death or a prolonged disorder of consciousness like vegetative or minimally conscious state. This prognosis informs treatment decisions about this patient, specifically the decision to withdraw life-sustaining treatment. When life-sustaining treatment is withdrawn based on the prediction of 'poor outcome', Chris dies as a result. This makes the prediction true regardless of whether continued treatment would have led to significant recovery. Thus, the prognosis is a self-fulfilling prophecy².

In what follows, I will first lay out the descriptive account of SFPs which forms the essential basis for understanding the normative analysis (§2). I will then proceed with the normative analysis which illuminates both the epistemic issues SFPs pose as well as their moral relevance (§3).

2. DESCRIPTIVE ACCOUNT: WHAT ARE SFPs AND HOW DO THEY WORK?

In order for a prediction to become self-fulfilling, several conditions must be met. Below, I first list the four required steps for a prediction to be self-fulfilling; credibility, employment, employment sensitivity, and realization (§2.1 - §2.4). Together they form the basis of the definition of an SFP (§2.5). Finally, I elaborate on the notion of collective SFPs (§2.6).

2.1. Credibility as a prediction

In order to become actionable, a prediction must be relied upon. This means that it must first be given credibility. A prediction can be given credibility for many different reasons.

Suppose a family gives credibility to the prognosis simply because they treat the medical professional with deference due to their trustworthy or authoritative status. Any prediction issued by the professional is then treated with *deferential credibility*, even though the professional may themselves grant credibility to the prognosis for a variety of reasons.

In chapter 4, for example, I described an instance of predictive infrastructure where the monitor reveals to the practitioners the predicted outcome, circumventing the need for the medical professional to understand the underlying patterns that led to the prediction. The practitioner can thus grant *default credibility* to the prediction which informs, without further scrutiny, the next step of processing or behavior. Especially with automated prognostic tools, depending on how they are designed and implemented, default credibility may be the norm.

² This SFP has been discussed in bioethics literature (Wilkinson 2009, 2013) and in debates about the quality of medical research (Geocadin et al. 2019)

In contrast, when professionals analyse the patterns themselves manually, they nevertheless rely on a theory or model that is already treated as credible. For instance, if they rely on the predictive model that an iso-electric pattern is predictive of a poor outcome, the prognosis is given *theoretical credibility*.

Family members and practitioners may give credibility to a prediction of poor outcome, even if they believe that there remains a degree of uncertainty, simply out of fear for the undesirable consequences that such outcome can entail. *Fear-induced credibility* can be granted for the sake of forestalling these consequences.

In contrast, consider an instance where the prognosis predicts full neurological recovery. It may be that this prediction is treated as credible, regardless of remaining uncertainty, simply because it is agreeable. *Agreeable credibility* can be given due to a combination of moderate plausibility and something akin to wishful thinking.

As we will see later on, collective SFPs typically are subject to *herd credibility*. This occurs when a prediction is given credibility because others give it credibility.

Placebo-effect is a typical example of a prediction that can be given *employment-dependent credibility*. There is an expectation that predicting (here, the efficacy of treatment) and acting upon that prediction will have a self-fulfilling effect. For that reason alone, the prediction can be given credibility.

The examples listed here comprise a non-exhaustive list of *sources of credibility*. Credibility can be granted for other reasons. Furthermore, it is common for several sources of credibility to coincide. A family's granting credibility to a prognosis may be the result of a combination of deferential, theoretical, fear-induced, and herd credibility. Granting credibility to the prediction that a sugar-pill may have a curing effect, can be a combination of theoretical, agreeable, and employment-dependent credibility.

This variety in potential sources of credibility distinguishes practical prediction from prediction in its main scientific role. In prognostic research, for instance, a prediction is granted credibility simply when it is entailed by a theory that is under consideration. An example is the theory that cEEG measuring brain activity patterns over time can project how much brain damage a patient has suffered. The credibility granted here is a warrant for a narrow range of further actions that are part of research and development. Namely, observing whether the outcome was as predicted, in order to further confirm or disconfirm the theory.

When analyzing the credibility given to predictions, it is important to note that mere assertions, when facing varying degrees of uncertainty, can be given credibility

as a prediction too. Therefore, my account requires a broad notion of 'prediction'. The essence is that a degree of uncertainty is met with a degree of assertoric force. Predictions, like assertions, but unlike suppositions, questions, speculations, etc. are intended to allow the hearer or recipient of the prediction to rely upon it as true, or at least actionable (Bach et al.1979). Thus I define a *prediction* as a record (a momentary or persistent information bearing state) of a proposition about a future observation or other yet unknown state of affairs, which is presented with something akin to assertoric force.

To make this more tangible, consider the normative uncertainty regarding what physiological outcomes constitute a 'poor outcome', discussed in chapter 3. Now consider the assertion 'severe neurological damage constitutes a poor outcome'. This is a signal, intended to be informative, about a proposition that hitherto had been regarded as less certain. Crucially, such signals can be given when the receiver, but often the sender too, is unaware of or uncaredful about the uncertainty that is really at stake. This means that no one may in fact realize that a prediction was made. Whether or not an assertion is recognized as a prediction, if it is given credibility as a prediction, it becomes actionable as a prediction.

2.2. Employment of the prediction

Credibility makes a prediction eligible for employment but giving credibility to a prediction does not guarantee that that prediction will be used. Credibility as a condition is thus insufficient. As soon as the prediction is relied upon, however, we can say that the prediction is employed.

There are many explicit ways in which a prediction can be employed. For instance, once the intensivist relies upon the prognostic test, that prognosis is used. Even if the prediction is not employed to inform decisions on whether or not to withdraw life-sustaining treatment. The intensivist can use the prognostic data to inform other treatment decisions without ever communicating the information itself. Or, more likely, they can employ the prediction merely by informing the patient's family about what outcome they can expect. The family, in turn, can rely on the intensivist's assertion about the outcome of the patient. They may employ the prediction to consider the wishes of the patient and inform the intensivist. Finally, the prediction can be employed to inform decisions on withdrawal of life-sustaining treatment.

The employment of a prediction thus consists of the actions and operations performed on the basis of the prediction. Typically, as in the example of a prognostic test, these actions or operations constitute the immediate, if not the ultimate, purpose for which the prediction was produced.

While explicit linguistic expression of the prediction is one paradigmatic example of the ways in which predictions are typically employed, it doesn't always have to be an oral declaration such as the doctor informing the family. Since linguistic expression can also happen through display on a screen or publishing in some other way, these expressions can happen automatically. As illustrated in chapter 4, one result of such automation is that medical professionals and sometimes family too can see the prognostic information before the neurologist officially informs them.

Perhaps the most common way for humans to employ a prediction is to form an expectation of, and often an opinion about, the predicted outcome. When a prognosis is communicated and given credibility, it automatically yields an expectation of that outcome. That expectation typically comes with a value judgment or, at the very least, questions on how to value the prediction. In standard human contexts, these simple employments of the prediction thus follow immediately on the heels of credibility. Cognitively treating a prediction as credible, immediately yields a belief of the prediction and an expectation of that realization. The computational equivalent is simply storing the predicted value for subsequent reference which usually also entails some way of categorisation. Take, for instance, the automated generation of a predictive value through cEEG, which is simultaneously categorised as good, poor, or uncertain.

Value judgements and categorizations are typical employments where input has been used for further inference or processing. These are relatively simple deductive inferences. More complex processing like predictive modeling where the content of the present prediction serves as an input for another prediction task are common in medical prognosis. In chapter 7, I further elaborate on how exactly these processes can lead to and exacerbate self-fulfilling prophecies and their moral implications. Employment of the prediction is absolutely essential. Without it, a prediction is inert and can never be self-fulfilling. If we think of credibility as opening a door, employment is pushing the prediction across the threshold.

2.3. Employment Sensitivity

On the other side of the threshold lies the object the prediction is about. That object can be sensitive to the employment of prediction. SFPs are necessarily *sensitive* to the prediction's employment. To put it in the most basic terms, in order to self-fulfill, the object of prediction has to somehow be sensitive to the prediction. Only when it is sensitive to the prediction's employment can the prediction itself have an impact. For that impact to be self-fulfilling, the object the prediction is about has to have either interpretation sensitivity leading to interpretative fulfillment or substance sensitivity leading to substantive fulfillment. I will demonstrate the difference between the two *sources of fulfillment* by using examples of placebo effect.

Consider a patient who has been suffering acid reflux, chest pain, and shortness of breath. Assume that the prediction 'this pill will make the patient better' is given credibility by the patient even if the medicine in question is just a sugar pill which has no inherent medicinal effect. The patient employs the prediction by taking the pill. Consecutively, the patient starts feeling better. Their symptoms disappear completely. This is a typical scenario of placebo-effect in medicine observed in most clinical trials. Now consider two variations of this scenario.

First, let's assume that the patient's symptoms were initially diagnosed to be the result of heartburn caused by a hiatal hernia, meaning the upper part of the patient's stomach bulges through their diaphragm into their chest cavity (Mayo Clinic, 2020). In one scenario, after the patient takes the placebo pill and their symptoms disappear, the patient's diagnosis is checked again and the hiatal hernia can no longer be detected. This would mean that the fulfillment was substantive. The actual substance the prediction was about transformed to match the prediction. In this particular case, it is difficult to explain what causes the substance sensitivity because it is not always clear what causes a hiatal hernia. It may be that the promise of getting better reduced the patient's stress to the point where they physically relaxed enough to reduce the persistent and intense pressure on the surrounding muscles that were causing the hernia. In any case, the self-fulfillment is due to *substance sensitivity*.

Now consider an alternative scenario in which a second diagnosis, post-placebo, reveals that the hiatal hernia is still present. Moreover, further examination shows that the hernia still causes acid reflux. Yet, the patient no longer experiences chest pain or shortness of breath and doesn't experience the acid reflux as cumbersome as the way they used to. The patient's being 'better' here is due to interpretative fulfillment. It is the interpretation of the substance the prediction was about, not the substance itself, that made the patient better. This fulfillment is thus due to *interpretation sensitivity*.

As this example shows, whether interpretation sensitivity can be the source of fulfillment depends on how the prediction is formulated. Specifically, it depends on whether the formulated object of prediction leaves room for interpretation. If the original assertion had been 'this pill will cure the patient', would that prediction have fulfilled itself? Much depends on what we consider 'curing'. If curing entails the patient's lack of experienced symptoms, then the answer is yes but the ways in which the patient experiences symptoms is an *observer-dependent* property. In contrast, we may expect that the examination of the patient's physical state and the diagnosis of hiatal hernia consists of objective measures that remain identical over time. Hence, both the evaluations before and after the placebo was administered are subject to the same standards, leaving little room for subjective interpretation. As

such, these measures are not as observer-dependent. Ultimately, the more specific the description of a prediction is, the smaller the room for interpretation will be. If the prediction is 'taking this sugar pill will reduce the patient's stress, restore the natural position of their stomach, fix the hiatal hernia, and subsequently, make all symptoms disappear', the room for interpretation will be significantly reduced, if not eradicated. This, however, does not prevent an SFP from occurring. The prediction could still be self-fulfilling due to substance-sensitivity.

Finally, it is preferable to speak of a *system* rather than an object being sensitive to employment. This way one can encompass a broad range of interactions between the predictions and the predicted, and the diverse factors that may mediate those interactions. In the next chapter, I demonstrate how the assertion that specific physiological outcomes are either good or poor can become self-fulfilling due to either interpretative or substantive fulfillment. It will become increasingly clear that SFPs do not always unfold by the prediction's employment operating directly upon the objects or characteristics mentioned in the prediction. An assertion that severe neurological damage entails a poor outcome, may affect the valuation of that physiological outcome in a roundabout, indirect way. Focusing on systems allows us to acknowledge a wide range of mediating factors and indirect effects.

2.4. Realization of the prediction

Regardless of the source of fulfillment, whether SFPs are fulfilled due to substance or interpretation sensitivity, the actual realization of the SFP can be operative or transformative. As I stated at the beginning of this chapter, there is a fundamental misconception that SFPs are relevant only when substantive or interpretative fulfillment impacts the actual outcome so that it turns out different from what the outcome would have been, had no prediction been made.

Let us return to Chris' case, where a poor prognosis led to the withdrawal of life-sustaining treatment. This employment of the prediction directly caused his death whether or not he would have died otherwise. To see the significance of this, compare the following two hypothetical cases.

Imagine Abe and Bill. Like with Chris, their prognosis indicates a poor outcome. Based on those predictions, life sustaining treatment is withdrawn, in effect causing them to pass away. For both of them, it was the prediction that led to their deaths. Let us assume, however, that in Abe's case, the prediction did not change the outcome. The prognosis was based on valid information and Abe would have died anyway. Here, the employment didn't change the outcome of the patient but it did affect the way the outcome came about. For Abe, the employment of the prediction was operative in bringing about the outcome, even if the same outcome would have

been realized without the prediction and its employment. This is an *operative self-fulfilling prophecy*³.

Now, let us assume that the prognosis Bill received was based on an incorrect understanding of his condition. Bill would have actually survived without any severe disability had life-sustaining treatment been continued. Now the SFP did not only change the mechanism of death; it caused the death of a patient that would have otherwise lived. It flipped the outcome from good to poor. In such a case, the predicted outcome is not merely *explained by* the prediction; the predicted outcome *depends* on the prediction. An SFP that flips the outcome to match the prediction, I call a *transformative self-fulfilling prophecy*.

Clearly, transformative SFPs demand our immediate concern in ways that operative ones do not. However, not recognizing an operative SFP as a genuine SFP is the same as not acknowledging that the genuine cause of a boxer's death is the accidental lethal blow of his opponent, just because—counterfactually—the boxer would have collapsed from heart failure soon thereafter. Furthermore, recognizing both as genuine SFPs is important because, in practice, it is not self-evident how to distinguish between the two. In fact, it may be impossible to find out, in retrospect, whether the SFP did or did not change the outcome. In neuroprognosis, having no complete certitude about the eventual fate of the patient is precisely the problem. Finding out, in retrospect, what *could have been* is impossible. As we will see in the normative analysis that follows the descriptive account, this inability to distinguish between operative and transformative SFPs lies at the heart of the problems that SFPs pose.

Finally, it is important to note that predictions, especially optimistic ones, can be given credibility and be employed, but while the object of prediction is—at times—sensitive to the employment, the employment does not end up realizing the prediction because the employment sensitivity requirement in that particular instance failed. Because they do not self-fulfill, they are not SFPs. As we will see, these predictions are nevertheless relevant to the normative analysis.

³ Wilkinson (2009) calls this a 'self-reinforcing prophecy'. I opt for the label 'operative self-fulfilling prophecy' for two reasons. First, since the employment of the prediction actually explains the outcome predicted, it is a genuine SFP, not merely a reinforcement of a prophecy. Second, 'reinforcement' does not signal any change of mechanism, despite the fact that, in such cases, the would-be means by which the outcome would be realised is *displaced*, not reinforced. In an operative SFP, the prediction and its employment become operative, in place of the mechanism by which the prediction would have been realized.

2.5. Defining the self-fulfilling prophecy

The 4 steps discussed in the previous section form the conditions that are individually necessary and jointly sufficient for a prediction to be self fulfilling.

#1 Credibility as a prediction

The prediction is treated as credible by the parties that employ the prediction.

#2 Employment of the prediction

Some action or operation constitutes the employment of the prediction

#3 Employment sensitivity

The system the prediction is about, in its substance or interpretation, is sensitive to the employment of the prediction.

#4 Actual realization

The prediction is actually realized due to the system the prediction is about being sensitive to and affected by the prediction's employment.

In light of the account explaining the mechanism of an SFP, I define SFPs as follows:

A self-fulfilling prophecy is a prediction, treated as credible, that is realized due to the subject of prediction being situated in a system that is sensitive to and affected by the way the prediction has been employed.

As stated before, my account of prediction includes any situation where a degree of uncertainty is met with a degree of assertoric force.

While this definition of SFP is significantly different from the original one by Merton who coined the term in 1948, it offers several meaningful contributions required to understand the normative problems regarding SFPs. First, Merton's definition 'The self-fulfilling prophecy is, in the beginning, a *false* definition of the situation evoking a new behavior which makes the originally false conception come true' (1948) applies only to transformative SFPs and is therefore incomplete. Second, excluding operative SFPs in that way hides an important characteristic of SFPs. Namely, that it is difficult, and sometimes impossible, to distinguish between transformative and operative SFPs. This means that the limitations of Merton's account would apply not only to transformative SFPs but *only* to those we can recognize as transformative, or *potentially* transformative. This is unacceptable. As we shall see, these kinds of epistemic problems are precisely at the heart of the normative relevance of SFPs that has been missed in the past. Second, this definition diverges from that of Wilkinson, even though Wilkinson does in fact acknowledge many of the epistemic challenges that SFPs pose (2013).

I argue that the account and definition of the SFP given here corrects mistakes of the past, thereby allowing us to either highlight the moral relevance that was

previously missed or clarify epistemic issues that have been identified but poorly understood. In what follows, I will offer a normative analysis of the practical and epistemic issues SFPs pose and how one fails to take responsibility for them. Before I get to that, I will expand the descriptive account with a short elaboration on collective SFPs.

2.6. Collective SFPs

This account, so far, describes individual instances of SFP. Yet, there are also cases where many predictions, or many employments of the same prediction lead to self-fulfillment. And, as we shall see in chapter 7, such *collective self-fulfilling prophecies* are at times very relevant to medical research as well as clinical practice. In light of this, I now offer a modicum on *collections of employments* that are *together* self-fulfilling.

One of the most common examples of collective SFPs in medicine, although they have not as of yet been recognized as such, relates to the medicalization of specific 'conditions' and their public evaluation. Typical examples are deafness, dwarfism, and Down syndrome. Medicalization is the perception and treatment of these conditions as health defects or disabilities, often implying that they require fixing or eradication (Conrad 1992, White 2017).

When a specific condition is asserted as being a disability that requires treatment, and that assertion is given credibility, it is much more likely that treatment will be offered and accepted. As such, the treatment of the condition equals the employment of the prediction. The effect of many such predictions receiving credibility and subsequently being employed, is that an increasingly smaller population of people with this condition remains. Unfortunately, this may also create troublesome effects of further alienation of, and pressure to assimilate for, those that remain. As such, medicalization has been understood as a form of social control (Zola, 1972, White 2017). Moreover, and most relevant to our discussion, the effects of a reduced population can further reinforce the assertion that the condition is indeed a disability that needs fixing and eradicating, thereby fulfilling the original prediction for those who had not given it credibility before.

The sources of self-fulfillment in these scenarios can be interpretative as well as substantive depending on what happens as a result of the population reduction. Take, for instance, in the case of deafness, treatment by way of cochlear implants. By systematically fixing deafness, its prevalence is reduced. Perhaps that is the right course of action. We do not mean to make a value judgment here. However, the confrontation and experience of the majority (here the hearing population) with the condition (here deafness) is increasingly limited and, as a result, the minority (here the deaf community) is increasingly misunderstood. While deafness, to

those who live with it, is in no way more or less of a disability than before the prediction was made, the hearing community increasingly interprets it as one. This is interpretative fulfillment. However, majority rule tends to also have substantive effects on society, its structure, its organisation, and its services. When a reduced population leads to a reduction in schools, transmission of sign language, or general exclusion, deafness will, in effect, become more debilitating. The employment of the majority's prediction can thus create an environment in which the minority is increasingly experiencing their condition as a limitation rather than a mere trait. This is substantive fulfillment.

Whether interpretative or substantive, collective SFPs thus require a collection of employments. In such cases, the prediction, once given credibility, can prompt the decision to treat a single occurrence of deafness with a cochlear implant. But no single instance of such employment of the prediction is sufficient for self-fulfillment. The system is only sensitive to the combined effect of many such employments. These employments can be the result of a singular prediction, given credibility by many, or multiple predictions predicting the same outcome. Either way, it is the collection of employments that are together self-fulfilling.

What is fascinating about these types of collective SFPs, is that they work in both directions and can thus often be reversed if so desired. In many societies, for instance, sign language use is mostly limited to deaf people. Typically, hearing people in such societies perceive being deaf as a serious problem requiring professional intervention (Lane 2005). This often leads to self-fulfilling feedback such as in the example above. In contrast, societies like those on Bali and Martha's Vineyard, have adopted widespread use of sign language because of a substantial deaf population. As a result, being deaf is regarded as a trait there, not a disability (Lane et al. 2000). This example gives us an idea of how such collective SFPs can be reversed. Imagine we assert that deafness is a trait, perhaps to some unfamiliar but a valuable trait nonetheless, and as a result sign language is taught in schools for the hearing. Suppose it is offered alongside other foreign languages and students could just pick among them. This would give self-fulfilling feedback in the other direction, asserting that deafness is rather exotic but not something that requires fixing.

Although I highly encourage further study of collective SFPs, I limit my account here to what is needed for illuminating the analysis in chapter 7. I now turn to the normative analysis of self-fulfilling prophecies, based on the descriptive account above.

3. NORMATIVE ANALYSIS: WHAT IS PROBLEMATIC ABOUT SFPs?

Now that we clearly understand how SFPs work, we can better analyze what is problematic about them. The descriptive account has serious implications for the ways in which one must take practical and epistemic responsibility for the consequences of SFPs. I will first discuss the existing concern (§3.1.) and then correct the epistemic critique I find inadequate (§3.2).

3.1. Current concerns regarding unintended outcomes and accountability

The first reason that distinguishing between operative and transformative SFPs is so important is that the focus of practitioners as well as critics has exclusively been about transformative SFPs⁴ (Gold et al. 2014, Amorim et al. 2018, Hofmeijer et al. 2019, Geocadin et al. 2019, Endisch et al. 2020) and we can see why. As explained at the beginning of this chapter, they have much more immediate practical consequences that require our attention. Clearly, the difference between causing an inevitable death (as with an operative SFP) or one that could have been prevented (as with a transformative one) is significant.

It is therefore no surprise that the focus on transformative SFPs, especially with pessimistic ones, has also led to a concern that SFPs help evade responsibility in an important way, regarding the practical consequences of the outcome. If we only look at instances where the prediction was not only responsible for the outcome but the outcome *depended* on the prediction and would not have been realized otherwise, where does that leave us with regard to accountability? Clearly the predictor, or those who gave credibility to or employed the prediction are responsible for bringing about the outcome.

The existing concern is that SFPs evade accountability because of some epistemic failing. The epistemic failing, intuitively, seems to be about some unjustified belief about what brought about the realization of the prediction. Surely, the unjustified belief cannot be about giving credibility to the prediction, since the prediction ends up being accurate. This is not a question of accuracy. The fallacy is a mistaken belief about what (or who) makes the prediction accurate (Biggs 2009, Dowden 2020). Crucially, it is about being unaware of the role one played in bringing about the outcome. As Biggs puts it, 'The actors within the process—or at least some of them—fail to understand how their own belief has helped to construct that reality' (2009).

⁴ The focus on transformative SFPs is such that the term SFP tends to refer exclusively to transformative ones (e.g. in the sources referenced).

The argument regarding responsibility that follows is that if there is total unawareness, then one cannot be answerable for producing the outcome and therefore one cannot be held accountable. And thus SFPs, by cloaking the agents' contribution, help evade accountability for those who help bring about the outcome. I argue, however, that there are some serious problems with this argument.

First of all, unaccountability due to unawareness in the past does not entail unaccountability for future instances. I would argue that *if* there's total unawareness *and if* you cannot be blamed for that unawareness, then you can indeed not be held accountable. Unawareness is only an excuse for unaccountability as long as you are not to blame for the unawareness. This means that, even if you were not responsible for a mistake in the past, once there is awareness about the mistake, you may be held accountable for reproducing the mistake in the future. Clearly, if the problem is a lack of awareness then the solution is an increase of awareness. I will return to the importance of increased awareness later since it may not be significant here, as my next point will prove.

Second, (un)awareness may not be as relevant as we think it is when it comes to practical prediction. We can see this in medical prognosis, early-diagnostics, and clinical research where there often *is no* unawareness. Let us look at cases where there is explicit acknowledgment of the potential for transformative SFPs and their consequences.

For those that have a positive effect, like placebo, the awareness will obviously not give reason to prevent employment of the prediction. On the contrary, if the prediction self-fulfills, great. But even explicit acknowledgment of pessimistic SFPs potentially being transformative may not prevent an agent from employing the prediction. Take transformative SFPs for coma patients that result in a poor outcome when a poor outcome could have been prevented. Once aware of this possibility, different practitioners make different decisions while taking this risk into account. Some practitioners may indeed choose to avoid employment of the prediction altogether (Gold et al. 2014). They may choose not to withdraw life sustaining treatment. But this may come at a significant cost to the patient as well as to the distribution of resources to other patients. In essence, whether the prediction is employed and how, is often based on a mixture of time-sensitivity, balancing costs and benefits and, ultimately, it is a judgment call. Even if the practitioner is responsible, they may still just be doing the best they can, precisely because there is so much uncertainty. Of course, uncertainty is inherent to prediction. Without it, there would be no need to predict.

Medical decision-making in general and intensive care decisions especially are marred by uncertainty. For instance, think about an unconscious patient who

likely has condition X. I say 'likely' because the patient shows symptoms of X and the intensivist, based on their experience, can recognize it as a potential case of X. Yet, when a scan is made to test the presence of X, it does not show up on the scan. This, as intensivists confirm, does not mean that X is not there, it just means they cannot find it. Now imagine that, to the best of the intensivist's knowledge, treatment Y is likely to improve the patient's condition. As a result, the intensivist predicts that the patient will benefit from treatment Y. However, there is a small chance that the patient is allergic to treatment Y. If the patient is indeed allergic to treatment Y, this would cause massive complications in their condition. Now, the intensivist can test the patient for allergies but receiving test results would require a wait of three hours and the patient will most benefit from the treatment if they get it as soon as possible and at the very latest within the next hour.

To be clear, this is a regular prediction, or asserted educated guess, made in light of uncertainties. The intensivist is fully aware of the risk they are taking by administering the treatment. Is the intensivist responsible for the complications when it turns out the patient is indeed allergic to the treatment? In some sense, sure. But holding them accountable for the consequences of their decision seems besides the point. In practical prediction, predictions generally take place in situations where decision and action is desired, despite deficient knowledge of all the relevant factors. Clearly, being aware about the effect of one's actions does not solve the problem. Therefore, unawareness is not the issue. When uncertainty faces time-pressure, cost and benefits must be weighed and a judgment call must be made, often prematurely. There simply may be no better way of responding. Administering the treatment may be the most responsible thing to do, *even* if it later turns out that it was a mistake.

When it comes to SFPs, this last bit will prove to be crucial; '*if it later turns out that it was a mistake*'. I will now correct for what I believe to be an inadequate epistemic analysis of SFPs.

3.2. Taking practical and epistemic responsibility for self-fulfilling prophecies

The big difference between other, regular predictions and self-fulfilling prophecies is that with SFPs, if a similar error was made, and the initial judgment call was flawed, it will not later *turn out* that it was a mistake because the SFP will prove the prediction right *even* if the initial judgment call was flawed.

Think of the hypothetical cases of Abe and Bill, discussed in the descriptive part of the chapter. Bill would have supposedly survived had life sustaining treatment been continued, hence making it a transformative SFP. But how would we know? In real life, once the prediction is employed by ways of withdrawal of life-sustaining

treatment, the patient dies and we cannot find out in retrospect what would have happened instead. Unlike in the case of treatment Y, where we see complications and allergic responses as clear error signals, Bill's death produces no such error signals. So we do not find out that it was a mistake. What's worse, when Abe passes away after withdrawal of life-sustaining treatment, even if the judgment call there was correct and Abe would have passed away anyhow, we cannot find out that the judgment call was right either. The lack of error signals here is not a reliable sign of positive feedback.

SFPs, because they entail their own fulfillment, do not produce reliable feedback. And if we cannot accurately identify our failures or our successes, we cannot accurately learn from them either. This is the central epistemic failing of SFPs.

Typically, error signals indicate when a mistake was made and a lack of error signals can be interpreted as though the correct decision was made. We can learn from that. By never producing error signals, the feedback produced by SFPs which is *always* positive, cannot be relied upon. And so, not only can we not learn from our mistakes, we cannot learn from our successes either.

We can see now why it is very difficult, in practice, to differentiate between operative and transformative SFPs. The difference between an operative and a transformative SFP always depends on what would have happened instead, if no prediction had taken place, and knowledge about such counterfactuals is difficult. Not knowing when the prediction changed the outcome and when it did not is precisely why we cannot learn.

While this is equally true for optimistic as well as pessimistic SFPs, it may seem most urgent that we pay attention to learning about the pessimistic ones because their practical consequences are dramatic while those of optimistic SFPs are desirable. Yet, when it comes to learning, optimistic SFPs may be even harder to learn from.

I mentioned, in the descriptive account, the normative importance of predictions that are potentially subject to self-fulfillment due to the system—at times—being sensitive to the employment but end up not self-fulfilling because the employment sensitivity requirement *in that particular instance* failed. It's important here to distinguish between employments of optimistic and pessimistic predictions.

Consider a patient in coma with a poor prognosis whose ventilator is withdrawn at a time where patients are not expected to be able to breathe without the ventilator's help. The ventilator is removed, motivated by the expectation that the patient will pass away soon thereafter but the patient nevertheless starts breathing on their own. Unexpectedly, the system was not sensitive to the employment of the prediction. The prediction does not self-fulfill in which case we receive an error signal.

While employment of a pessimistic prediction typically leads to limiting options thereby reducing possibilities, employment of optimistic predictions tends to increase options, thereby increasing possibilities. With optimistic predictions, the system, even if potentially sensitive, is given much more room to respond to the employment in a variety of ways.

We see this with the placebo effect. The prediction 'this pill will make the patient better' when employed by communicating it to the patient and subsequently by administering a sugar pill, can have a transformative self-fulfilling effect making the patient better when they wouldn't have recovered otherwise. The employment can also have an operative effect, making the patient better even if they would have gotten better anyway. Finally, the employment can have no effect at all, thus not making it an SFP. Either the patient does not get better, in which case we receive an error signal or, the employment can have no effect at all but will still *appear* to be self-fulfilling, because the patient got better anyway for other reasons unknown to us. In this last case, even though the prediction does not self-fulfill, it appears as though it did.

As mentioned, the inability to distinguish between operative and transformative SFPs lies at the heart of the epistemic problems that SFPs pose. Optimistic predictions⁵ that end up not self-fulfilling but appear as though they do, muddy the waters further by making it hard to assess whether a transformative SFP, an operative SFP, or no SFP at all was at work. Optimistic self-fulfilling prophecies are thus even harder to clearly identify as such.

If the epistemic problem with SFPs is that they inhibit learning, where does that leave us with regard to responsibility? The general concern that 'SFPs evade accountability in an important way due to an epistemic failing' still rings true. The epistemic failing being, however, that SFPs inhibit learning. Therefore, the problem with SFPs is not that they help evade responsibility for the practical outcomes they create. Rather, they allow for agents to evade responsibility about their failure to learn. Perhaps some might think this is not a great moral injustice. After all, are we obligated to learn⁶?

⁵ Pessimistic predictions can in principle be subject to the same problem but it's much more unlikely as you would need the failure of the employment to go by unnoticed (without producing error signals) *and* simultaneously have something else, unrelated, happen that realizes the pessimistic prediction.

⁶ Some scholars in philosophy would argue that yes, we have an obligation to have adequate knowledge of the world. This obligation to 'know well' is what philosophers have termed 'epistemic responsibility.' (Code 2020, Corlett 2008) We can leave this debate aside here as there are specific reasons why our obligation to learn, and learn accurately, is important in this context. I lay out the arguments in what follows.

First of all, agents are likely to learn anyhow. Unconscious cognitive processes that automatically incorporate and organize information perceived as feedback lead to all sorts of unconscious—or implicit—learning (Kihlstrom 2009, Simons 2012). When relying on unreliable feedback, because one is unaware of the fact that it is unreliable, this will likely result in flawed learning processes. Specifically, one can learn to reproduce mistakes as though they were correct. Second, in medicine, there is at least one area where we expect practitioners to learn correctly and that is in research and development. Especially in prognosis, because there is so much uncertainty, researchers are constantly seeking to improve the ability to accurately predict. How is one supposed to improve a practice when mistakes and successes cannot be recognized as such?

I argue that correct learning is the most essential requirement of innovation. Responsible innovation therefore requires a critical approach to one's own learning processes. One has to know *when* one is learning, *what* one is learning, and *how* one is learning.

Here, finally, awareness *is* of crucial importance. If an agent is unaware of how SFPs are skewing their learning process—and—the agent is not to blame for that unawareness, then they cannot be held accountable. But, as I stated before, not being blameworthy for mistakes in the past, does not entail being unblameworthy for future instances.

The relevance of increased awareness about our failure to learn is not unlike that of increased awareness with regard to implicit bias. Agents may have unconscious cognitive processes that result in 'automatic associations, of which [they] may not be aware, that are difficult to control, and may conflict with [their] professed beliefs' (Holroyd et al. 2017). It may be difficult to hold an agent accountable for the unintended, practical consequences of their judgments and actions caused by such implicit biases. 'The[se agents] may, however, be blamed if they fail to act properly on the knowledge that they are likely to be biased e.g. by investigating and implementing remedies to deal with their biases' (Saul 2013). I argue that, similarly, agents can be held accountable if they fail to act properly on the knowledge that they are likely to err in their learning processes.

Improving our ability to learn will require in depth study of the relationship between the employment of predictions and the sensitivity of the systems the predictions are about. Distinguishing between optimistic and pessimistic predictions proves to be crucial in that regard. While delving into this in further detail here will take us too far afield, it is important to note that the employments of optimistic versus pessimistic predictions differ greatly in how they affect employment sensitivity.

Specifically, optimistic employments⁷ tend to open up the ways in which systems can respond, thereby making the prediction less vulnerable to employment sensitivity, while pessimistic employments generally narrow the options down, making predictions more vulnerable to employment sensitivity. As a result, however, when optimistic predictions are realized, it is much harder to identify the source of realization and, thus, whether it was indeed the employment that caused or contributed to self-fulfillment.

With placebo, even if the employment sensitivity is not well-understood, the recognition of their self-fulfilling effect led to a new standard for testing our learning processes in clinical research. I am tempted to argue that the recognition of some pessimistic self-fulfilling prophecies require us to do the same. In the next chapter, I describe in much more detail the practical and epistemic consequences that result from SFPs in neuroprognostication.

3. CONCLUSION

This chapter presents a descriptive account and normative analysis of self-fulfilling prophecies in medicine. The main goal is to correct for shortcomings in existing accounts in order to better understand the problems SFPs entail in medicine and how to address them.

The common concern regarding SFPs is that they transform the predicted outcome to match the prediction. Such transformative SFPs indeed have some special moral implications for practical prediction. However, SFPs pose an epistemic challenge as well. To understand the core epistemic challenge, we have to analyse *all* SFPs, including operative ones that do not change the outcome but only impact the way the outcome was brought about.

The descriptive account offers a detailed description of the mechanism of SFPs and their different forms. The four elements, credibility, employment, employment sensitivity, and realization, together form the conditions that are individually necessary and jointly sufficient for a prediction to be self fulfilling. As such, they form the basis of the definition, 'a self-fulfilling prophecy is a prediction, treated as credible, that is realized due to the subject of prediction being situated in a system that is sensitive to and affected by the way the prediction has been employed'.

⁷ I speak here of optimistic and pessimistic **employments** as opposed to **predictions** for comprehensiveness. Generally, employments can be optimistic regardless of whether or not the prediction is optimistic (as when the prediction is not given credibility or is employed optimistically regardless of the credible prediction). Since these instances (where employment does not reflect the prediction) will never lead to SFPs however, they are outside of the scope of this article.

The normative analysis focuses on illuminating SFPs' core epistemic problem. SFPs inhibit appropriate learning due to their automatic confirmatory feedback. Whether or not one succeeded or failed to predict the eventual outcome absent of any prediction and its employment, one cannot find out. This inability to retrospectively check the validity of the prediction is an impediment for research and innovation. In medicine, this limitation can have dramatic consequences.

In the next chapter, I will analyze the epistemic relevance for neuroprognostic research and innovation and the ethical implications this has for clinical practice.

Chapter 7

Self-Fulfilling Prophecies & Responsible Neuroprognostic Innovation:

recognizing invisible physiological and normative prognostic errors

ABSTRACT

True predictions are typically hard to fault, but what about self-fulfilling prophecies (SFPs) that produces its own truth? In neuroprognostication, when predictions of patient outcome guide the decision whether to continue life-sustaining treatment, the stakes are life-and-death. Although the problem of false positives leading to SFPs in neuroprognostication is recognized in both the medical and bioethics literature, the prevailing conceptual analyses of SFPs and their epistemic as well as moral status are far from complete. Adequate evaluation of neuroprognostic practices requires us to, first, properly identify *all* SFPs, including those that do not change the outcome of the patient. Such evaluation entails a move away from focusing on false positive test results only. Second, we must identify two types of SFP in neuroprognostication; one in acute care when facing physiological uncertainty and one in chronic care when facing normative uncertainty. Finally, we must analyse the effects of physiological and normative SFPs at two interconnected levels in both acute and chronic care; first regarding the medical outcomes for particular patients and second regarding the consequences for members of the patient group. I argue that, aside from sometimes resulting in undesirable outcomes, SFPs inhibit learning. This makes decision-makers prone to repeat the same mistakes—both with regard to physiological and normative predictions of medical futility. Automated prediction through machine learning techniques may further exacerbate these problems. I show how SFPs of physiological and normative predictions, especially when iterated, endanger good medical practice and complicate innovation and policy in significant ways. Finally, I suggest interventions for identifying mistaken prognoses and breaking ensuing feedback loops.

The insights and arguments regarding physiological SFPs developed in this chapter have been published as **Mertens M, King OC, van Putten MJAM, and Boenink M (2021) Can we learn from hidden mistakes? Self-fulfilling Prophecy and Responsible Neuroprognostic Innovation. *Journal of Medical Ethics*. doi:10.1136/medethics-2020-106636**

Self-Fulfilling Prophecies & Responsible Neuroprognostic Innovation:

recognizing invisible physiological and normative prognostic errors

1. BACKGROUND: FILLING THE GAP

On Wednesday, Robert's heart stops beating. Although the emergency response team performs CPR successfully and restarts his heart, Robert does not regain consciousness. Hence, he is brought to intensive care in a postanoxic coma, or coma after cardiac arrest. Due to the lack of oxygen during the arrest, Robert is likely to have suffered brain damage. To reduce the damage as much as possible, the physicians are keeping him sedated while cooling his body to prevent fever. On Saturday, a test result measuring brain damage indicates Robert's prognosis is poor, meaning that he is predicted to have no chance of meaningful neurological recovery. According to this prediction, if Robert continues to be treated, he will either die or survive with severe neurological impairment. Based on the prediction, life support is withdrawn and Robert dies later that Saturday.

This scenario is common after out of hospital cardiac arrest (OHCA). Most postanoxic patients who die in intensive care later than 24 hours after the arrest, die after life sustaining care is withdrawn (Dragancea et al. 2017, Hasager et al. 2018, 989, 994). Hence, I focus on prognosis for postanoxic coma, precisely because patients who have been revived after cardiac arrest are patients for whom the prognosis of 'poor outcome' typically results in (early) withdrawal of life-sustaining treatment (WLST). This while (early) withdrawal practices tend to raise concerns due to their self-fulfilling effect.

In the previous chapter, I presented an empirically informed philosophical account of self-fulfilling prophecies (SFPs) in medicine in general¹, focusing on the conditions under which they take place and the moral consequences they entail. I asserted that the prevailing views of SFPs are incomplete because only one kind of self-fulfilling prophecy (SFP) is typically acknowledged as truly being self-fulfilling and deserving of our concern, those that change the outcome. Yet, following the discussion in chapter 6, the prediction that Robert would have a poor outcome, after given credibility, triggered a response that *guarantees* that Robert would have a poor outcome, regardless of how it would have turned out otherwise. In other

¹ The general account of SFPs in practical and automated prediction was developed in collaboration with Owen C King and has been submitted for publication as King OC and Mertens M (2021) Self-fulfilling Prophecy in Practical and Automated Prediction.

words, the prediction is a self-fulfilling prophecy whether or not the positive test result for brain damage was true or false.

The tendency for such poor prognoses of comatose patients in intensive care to be self-fulfilling has been recognized, and often criticized, in both the medical literature (Geocadin et al. 2019; Maciel et al. 2018, Amorim et al. 2018) and in bioethics (Wilkinson 2009, Geocadin et al. 2012) but most only recognize a *chance* of SFP after withdrawal. In this chapter, I show that when WLST leads to the death of the patient, it is *always* an SFP. By focusing only on the SFPs that change the outcome of the patient (or only recognizing those as true SFPs), no one has yet given a thorough and satisfactory conceptual analysis of SFPs in this context². Consequently, the moral evaluation of the practices and actions that give rise to SFPs in neuroprognostication is inadequate. This chapter attempts to correct both the conceptual as well as the normative deficit.

In this chapter, I will describe in detail the significance of SFPs to neuroprognostication for coma patients. My contribution differs from existing accounts in two significant ways.

- First, I will reemphasize that thorough evaluation of SFPs in neuroprognostic care and research practices requires us to acknowledge *all* SFPs, including those that do not alter the outcome of the patient, so that we can properly evaluate the ethical and epistemic implications.
- Second, a thorough analysis of SFPs in neuroprognostication furthermore requires identifying two types of SFPs (physiological and normative) and the different problems they pose at two interconnected levels, in the acute phase and in the chronic phase.

Earlier, in chapter three, I distinguished between physiological and normative uncertainty and argued that uncertainty regarding the two types of medical futility required making *two* predictions, rather than one. In this chapter, I show how not only predictions of physiological futility but also predictions of normative futility and utility can be self-fulfilling.

² While Wilkinson indicates that self-fulfilment is inevitable when life-sustaining treatment is withdrawn or withheld in the face of uncertainty, he refers to SFP as a self-reinforcing prophecy; a situation that merely increases the chance of the outcome to occur (2009). As mentioned in the previous chapter, the label 'self-reinforcing' is misleading for two reasons. First, since the employment of the prediction actually explains the outcome predicted, the prediction is responsible for fulfilment, not just reinforcement. Second, 'reinforcement' does not suggest any change of mechanism, even though, with an operative SFP, the means by which the outcome would have been realized is *displaced* or *replaced*, not reinforced.

The problem of *physiological SFPs* occurs in the acute phase, as in the example of Robert, when a positive test result for brain damage predicts a ‘poor outcome’ for a patient³. The resulting expectation is that this patient will either die, or survive with severe neurological damage or a disorder of consciousness. This prediction is then commonly employed in the decision to withdraw life-support, thereby ending the patient’s life. For these cases, the initial prediction caused its own fulfillment; namely, the patient’s outcome is poor regardless of whether it would have been poor otherwise. Importantly, when the prediction is also based on a mistaken prognosis, or a false positive as it is called in medicine, it is a mistake which goes undetected.

The problem of *normative SFPs* occurs in the chronic phase, after a prediction is given credibility that the lives of patients who survive with severe neurological damage will or will not be considered worth living and, as such, will constitute a good or poor outcome. Governments can, for instance, regulate end-of-life decision-making on that basis. These regulations affect the quality of life for these patients, further affecting whether or not their lives are worth living. Thus, the initial prediction can be partially responsible for which lives are worth living. This is a second, additional SFP. However, it has not been acknowledged as such. This chapter offers a first overview of the self-fulfilling mechanisms at work at this level.

I argue that these two self-fulfilling prophecies, each individually, endanger good medical practice regarding coma-patients who have suffered neurological damage. They furthermore complicate the decision-making regarding innovation and policy in significant ways. Moreover, they are interconnected and this interconnection makes each more complex. On the one hand, hidden errors impact individual patients and, on the other hand, these errors have consequences for members of the patient group which again loop back to predictions that impact individual patients. Nevertheless, since SFPs masquerade as ordinary true predictions, responsibility for the undesirable effects of these SFPs, if appropriate, is easily—even automatically—evaded.

In the next section, I first give a short summary of the key theoretical points introduced in chapter 6. I focus on the definition of SFP and the crucial distinction between transformative and operative SFPs. I then go on to thoroughly analyse, in five steps, the problem of physiological SFPs in acute care when predicting the neurological outcome of a patient in postanoxic coma. Section 3 then shows how SFPs’ inhibition to learning translates to three specific iterations in prognostic research and innovation. Section 4 introduces the normative SFPs that result from predicting what the normative evaluation of a specific neurological outcome will

³ In medicine, positive test results mean they indicate that the patient has the disease or condition tested for. This is counterintuitive, as it is understood as negative news. Vice versa, a negative test result indicates the patient does not have the disease or condition tested for.

be. I offer guidelines in section 5 by combining the relevance of both SFPs and suggesting several intervention points at which a change in medical and research practices can significantly reduce the false positives in physiological SFPs, the erroneous evaluations in normative SFPs, and break the ensuing feedback loops. I close with the conclusion, reemphasizing the upshot of the core challenge posed by SFPs.

2. FACING PHYSIOLOGICAL UNCERTAINTY. WHAT IS THE PROBLEM?

In the previous chapter, I defined an SFP as follows: A self-fulfilling prophecy is a prediction, treated as credible, that is realized due to the subject of prediction being situated in a system that is sensitive to and affected by the way the prediction is employed.

To make this tangible, imagine the prediction P 'this patient will have a poor outcome'. For this prediction to become an SFP, the prediction first has to be treated as credible by the doctors, who will then employ the prediction for decision-making on how to proceed with the patient's care. In this case, the employment is to withdraw life-sustaining treatment (see figure 7.1). The patient's physical condition is sensitive to this employment, in that their condition rapidly deteriorates without life support. And, thus, the actual realization of the predicted outcome is explained by the withdrawal of life support to which the patient's body is sensitive. The patient dies *due to* the employment of the prediction. The prognosis is self-fulfilling.

Figure 7.1. Pathway after prediction is employed by withdrawing life-sustaining treatment

Importantly, chapter 6 distinguished between *operative* SFPs where the employment of prediction brought about the outcome but the same outcome would have eventually realized even without the employment of the prediction, and *transformative* SFPs where the outcome *depended* on the prediction and its employment and would not have been realized otherwise.

Self-fulfilling prophecies, even those that bring about unfortunate outcomes such as the death of a patient, are thus not inherently problematic. When a patient is destined to die without ever regaining consciousness first, allowing them to pass away sooner rather than later may be a desirable course of action. One might argue that the suffering of the patient at the end of life is reduced, family is relieved from the prolonged uncertainty, and valuable resources are allocated to patients with chances of meaningful recovery. This is what happens in the case of operative SFPs.

Understandably, practitioners and researchers worry about cases of false positives for poor outcomes in which the patient whose life-support is withdrawn could have otherwise survived (Geocadin et al. 2012, Gold et al. 2014, Geocadin et al. 2019), thus transformative SFPs only. Yet, to fully grasp the ethical challenges that SFPs pose in neuroprognostic practice and research, we must analyse the different layers of a complex problem. Understanding what difficulty SFPs exactly add, will guide our understanding of how to responsibly respond to SFPs. In what follows, I carefully lay out the five layers of complexity that together explain a fundamental problem in neuroprognostication, caused by SFPs: that they inhibit learning.

2.1. First: remaining physiological uncertainty in neuroprognostication

If we were certain that the prognosis is absolutely reliable, SFPs would not be a problem. In other words, if the positive test result is always a *true* positive, all the SFP does is change the mechanism of death. The patient may still end up dying from heart failure, brain damage, or whatever is believed to cause the eventual poor outcome. However, that this happens as it does and *when* it does is due to WLST which takes over the natural process. This is thus a special kind of operative SFP where the prediction and its employment become operative in the progression and timeline of the disease. Withdrawal of life-sustaining treatment only causes the ultimate mechanism of death to be activated earlier. For cases like Robert's, that is often thought to be the desirable course of action.

The problem is that neuroprognostication is precarious. Typically, prognostication is highly uncertain (Smith et al. 2013) and, as long as there is some remaining uncertainty, errors can occur. When there is a chance that a positive test result is a *false* positive, the stakes are life and death. This is why the traditional worry in intensive care, regarding SFPs, is that a patient who could've recovered, now dies (Wilkinson 2009, Geocadin et al. 2012, Gold et al. 2014, Geocadin et al. 2019). Severe as this consequence may be, this patients' unnecessary death is not resulting from a moral issue regarding SFPs. In chapter 2, we saw that the root of this unfortunate, even tragic, outcome lies in a burdensome clinical dilemma caused by uncertainty.

2.2. Second: Probabilistic decisions are never error proof

As my fieldwork shows in chapter 2, healthcare professionals do not merely fear the death of a potentially healthy patient, they simultaneously worry about escalating medical treatment in patients whose health will be reduced below a desirable threshold. Their decision-making is pulled in opposite directions: On the one hand, practitioners aim to avoid premature withdrawal of life-sustaining treatment in patients with chances for recovery. On the other hand, caregivers aim to avoid continuing life-sustaining treatment in patients that will recover poorly.

Due to the uncertainty, it is often difficult to know which category the patient belongs to: the one with or without chances of meaningful recovery. The practitioners' objectives are thus in tension and the pursuit of either comes at a risk. Either erring on the side of withdrawal, with the risk that a patient who could have survived dies. Or, erring on the side of prolonging care, with the risk that a patient survives with severe neurological damage and disability.

When making probabilistic decisions on such precarious ground, physicians will inevitably sometimes make the wrong choice⁴. These mistakes are regrettable but they happen because of the dilemma caused by uncertainty. The intensity of this dilemma may furthermore be reduced or increased depending on the temporal pressure that forces practitioners to make decisions prematurely. Typically, vital automatic functions like circulation and breathing can return after two to three weeks of coma after which WLST may no longer lead to the death of the patient (Jennett 1975, Katz 2009). However, as discussed in chapter 2, different policies that allow withdrawal at different times may shorten or lengthen the so-called *window of opportunity*, during which withdrawal of life support will allow the patient to die (Kitzinger & Kitzinger 2013, Fisher & Ridley 2012). While the risk on one horn of the dilemma is due to the inability to detect a potential false positive *before the SFP* takes place, it is the pressure from the other horn of the dilemma that forces a premature decision, causing the risk of unnecessarily ending a patient's life. Once life sustaining treatment is withdrawn after a predicted poor outcome⁵ and the patient dies, only *then* is the prediction self-fulfilled. What problem does the SFP then *cause*? In the following two sections, we describe both an epistemic problem as well as a moral problem that ensues from it.

⁴ This happens for instance when life sustaining treatment is withdrawn after a(n undiscovered) false positive test result.

⁵ Decisions to withdraw life support do not require explicit predictions of poor outcome per se (e.g. when tests are not indicating a poor outcome but the patient's recovery nevertheless is perceived as unlikely). In that sense, a mere lack of good outcome prediction may be sufficient.

2.3. Third: SFPs produce no error signal nor reliable positive feedback

Uncertain decisions based on probabilistic diagnoses are common in medicine, and they can lead to fatal mistakes. These regrettable instances have nothing to do with self-fulfilling prophecies, however. The difference between a regular mistake made on probabilistic diagnoses and those that lead to a self-fulfilling prophecy is simply that the latter does not produce an error signal. In other words, because SFPs ensure their own truth, one will never be confronted with the fact that a mistake was made. Worse, due to the affirmative feedback, one may never even be called to investigate whether a mistake was made.

Usually, when doctors make decisions while facing uncertainty, they are afterwards confronted with the outcome. For instance, suppose that, based on a negative test result, a prediction (P-) is made that hypothetical patient Xander will have a good outcome and treatment is continued based on this prediction (see figure 7.2).

Figure 7.2. Pathways after negative test result

If the underlying medical assessment was somehow mistaken and the prognosis was based on a *false negative* (F-), Xander's outcome will be poor regardless of the continuation of treatment. The confrontation with this result teaches that a mistake was made. It becomes clear that the original prognosis was a false negative⁶. The error signal is apparent. Moreover, if the underlying medical assessment was correct and the prognosis was a *true negative* (T-), the patient survives with a good outcome. This positive feedback following a true negative contrasts nicely with the error signal following a false negative, making both sources of feedback reliable (see also table 7.1).

⁶ Even if there is no automatic appearance of a mistake, proactively checking whether a mistake was made will produce informative feedback.

NEGATIVE TEST RESULTS Outcome predicted: GOOD	
FALSE NEGATIVE	TRUE NEGATIVE
Life-sustaining treatment is continued based on good prognosis However, patient has poor outcome after continuation of treatment, based on the good prognosis (which is an error signal) Prognosis does not change outcome → No self-fulfilling prophecy	Life-sustaining treatment is continued based on good prognosis Indeed, patient has good outcome after continuation of treatment, based on the good prognosis (yielding no error signal) Prognosis does not change outcome → No self-fulfilling prophecy
Outcome observed: POOR → Observations provide reliable feedback	Outcome observed: GOOD → Observations provide reliable feedback

Table 7.1 - Feedback received when negative test results motivate continuation of treatment

Now compare this situation to Yann’s case, also hypothetical. Suppose that, based on a positive test result, a prediction ($P+$) is made that Yann will have a poor outcome and life-sustaining treatment is withdrawn based on this prediction (see figure 7.3).

Figure 7.3. Pathways after positive test result

If, once again, the underlying medical assessment was somehow mistaken and the prognosis was a *false positive* ($F+$), Yann will nevertheless die because of the choice that was made in response to the prediction. There is never a moment when it becomes evident that the prediction was a mistake. Because Yann’s prognosis, in addition to being a false positive, is *also* a self-fulfilling prophecy, the mistake is obscured; no error signal is produced. In addition, SFPs make affirmative feedback unreliable. Suppose that Yann’s prognosis was in fact a *true positive* ($T+$) and he would have passed away anyway or survived with severe neurological damage

had care been continued. Once life sustaining treatment is withdrawn, there is no reliable positive feedback that will reassure the physician that the test result was correct since the feedback is *always* affirmative (see also table 7.2).

POSITIVE TEST RESULTS Outcome predicted: POOR	
FALSE POSITIVE	TRUE POSITIVE
Patient dies after life-sustaining treatment is withdrawn, based on the poor prognosis However, patient would have had a good outcome (given continued life-sustaining treatment) Prognosis changes outcome → Transformative self-fulfilling prophecy	Patient dies after life-sustaining treatment is withdrawn, based on the poor prognosis Regardless, patient would have had a poor outcome (given continued life-sustaining treatment) Prognosis does <i>not</i> change outcome → Operative self-fulfilling prophecy
Outcome observed: POOR (death of the patient) → Observations provide unreliable affirmative feedback	

Table 7.2 - Feedback received when positive test results motivate withdrawal of treatment

While transformative SFPs (in cases of false positives) produce no error signals, operative SFPs (in cases of true positives) produce no reliable positive feedback. In sum, self-fulfilling prophecies do not produce informative feedback about whether or not credibility was justly given to the predictions. That is the epistemic problem that SFPs add to neuroprognosis.

By worrying only about false positives and thus only including transformative SFPs in their analysis, previous scholars have missed a significant part of SFPs epistemic and moral implications. It is an understandable mistake. The reason that medical and bioethics scholars have focused solely on transformative SFPs in neuroprognostication is that not knowing whether or not Yann dies in vain complicates clinical decision-making. Therefore, the issue regarding SFPs that physicians and researchers have been seeking to address lies in catching potential false positives that may go undetected *before WLST* and thus *avoid transformative SFPs* altogether. This is a natural response to the problem, since detecting a potential false positive *in time* is undeniably vital to the individual patient who may have a meaningful recovery if life-sustaining care is continued after a prediction of poor outcome turns out to be false. However, not acknowledging operative SFPs as genuine SFPs is the same as not acknowledging that the genuine cause of a boxer's death is the accidental lethal blow of his opponent, just because—counterfactually—the boxer would have collapsed from heart failure soon thereafter.

2.4. Fourth: Error signals and reliable feedback are fundamental to learning

Detecting a mistaken prognosis only matters to the individual patient at hand if you can detect it *before* treatment decisions are implemented and thus, before the SFP takes place. If an error signal can no longer save the patient, why care about error signals at all?

Even if the error itself cannot be corrected, we can still benefit from an error signal. Recall Xander who had a poor outcome even though his treatment was continued? Recognizing the false negative cannot make things any better for Xander, but if this mistake can be learned from, there is a higher chance of getting things right when encountering new patients Yolanda and Zachary in the future. So, *by learning from the mistake, one may redeem it* in at least some small way; managing to make Xander's tragic case a source of future good.

Now compare again with Yann who could have survived, had his life sustaining treatment not been withdrawn. Yann's death confirms the prognosis and thus, cannot produce an error signal and neither clinician nor researcher are ever in a position to notice their mistake. On the contrary, the affirmative feedback may seem to reinforce the prognosis.

Receiving an error signal after making a mistake is fundamental to every learning process. Unfortunately, our ability to learn from true and false negatives, for which we *do* receive reliable feedback and error signals, is not entirely independent from our ability to learn from true and false positives, for which we do *not* receive reliable feedback or error signals (see figure 7.4).

Figure 7.4. All pathways after any prediction is employed

In practice, this means that even correctly learning from true and false negatives may not be so self-evident. The prognostic capacities of the clinician treating Xander, even after realizing the prediction was a false negative, may not improve when deciding whether or not to withdraw life-sustaining treatment for Yolanda and Zachary. This is because a decision to withdraw, as long as uncertainty remains, would necessitate reliable positive feedback if we were to learn whether *that* was

the right choice instead. Not only do SFP's not produce error signals but they also do not produce positive reliable feedback since the positive feedback is due to self-fulfillment.

Rather than learning correctly, it is much more likely that the clinician, after sufficient false-negatives (and thus sufficient error signals), adjusts and chooses to withdraw more often than before, in an attempt to improve. But because the false-positive rate after withdrawal always appears to be 0%, the feedback for withdrawal is always positive. Even if the positive feedback cannot be trusted and clinicians should not conclude anything from the fact that Yann indeed died, they are still more likely to develop an evidential bias towards withdrawal in an attempt to learn from the error signals due to false-negatives. Due to the dilemma, the two learning opportunities are inherently connected.

Furthermore, this evidential asymmetry comes on top of an already existing asymmetry between the ways we evaluate harms and benefits. When faced with a choice between a risk of great harm or a loss of comparable or even greater benefit, precaution already creates a bias that favors avoidance of harm (Shiffrin, 2012). In medical practices where a choice has to be made between risking a terrible outcome, and risking the loss of a valuable one, this precautionary bias can be palpable⁷. It makes sense for clinicians and family members to favor avoiding the harm of subjecting a loved one to a severe disorder of consciousness or other quality of life they deem unacceptable. As such, this precautionary bias further favors the avoidance of false negatives, which compounds the evidential asymmetry between false positives and false negatives caused by SFPs.

2.5. Finally: Without correct learning we cannot accurately reduce uncertainty

To really solve the clinical moral dilemma, one must either take away the temporal pressure completely or eradicate uncertainty. Similarly, by reducing either uncertainty or temporal pressure one can also decrease the intensity of the dilemma.

The battle against uncertainty is the source that fuels the pursuit of better prognostic tools, like cEEG. The very goal of prognostic innovation is to increase certainty and thereby mitigate the severity of the clinical dilemma. Specifically, we want a high degree of certainty that, (1) when the prognosis is good, continued treatment will indeed yield a good outcome, and (2) when the prognosis is poor, continued treatment would not yield a good outcome.

The initial clinical dilemma already showed the difficulty in testing a new prognostic method for satisfaction of that second desideratum. Conducting

⁷ This need not always be the case. One can imagine that—with newborns in coma—the potential loss of an expected child is a harm that may flip the precautionary bias.

an appropriate experiment by continuing treatment carries an immense moral risk. At worst, it is unacceptably inhumane. If there is a credible prediction of a poor outcome, then the only straightforward way to validate this prediction is to continue treatment and observe whether this indeed yields a poor outcome (Wilkinson 2011). However, validating the prediction would entail bringing about exactly the sort of outcome, such as an unacceptable quality of life for the patient, that decision-makers aim to avoid. This means the clinical dilemma has an additional dimension, a choice between moral value in clinical practice and epistemic value in research: choosing either to prioritize the interests of the individual patient or to advance research for future patients to benefit from.

The other way, then, to address the clinical moral dilemma is to eradicate the temporal pressure. This is only possible by extending the window of opportunity past the point of uncertainty. In countries that have progressive regulations regarding withdrawal of treatment and end-of-life decisions (e.g. allowing discontinuation of intravenous nutrition and hydration⁸) continuing treatment to the point of certainty about the outcome is less risky. One has the option to withdraw care past that point, hard as that may be. Even then, however, a risk would still remain for a very small number of patients; those who would recover well physically (and thus fully regain consciousness) but be so cognitively damaged that they would be considered incompetent to make end-of-life decisions. If no prior arrangements are made, there is almost no way to undo the damage nor relieve the suffering of the patient and loved ones potentially caused by the continued treatment. Therefore, *some* uncertainty will remain, even with a very wide window of opportunity.

In short, ethical treatment of patients may *demand* the SFP, even though the SFP precludes the very observations that would be required to test the theories and technology on which the prediction was based. Unfortunately, when practitioners and innovators cannot learn from their mistakes, nor from their successes, attempts to innovate are subject to severe limitations—which I will now discuss.

3. LEARNING INHIBITORS ITERATE MISTAKES IN PREDICTIONS OF PHYSIOLOGICAL FUTILITY

In order to improve prognosis, new prognostic methods must be tested. Sometimes, research can be done outside of clinical practice but, for patients who suffered brain damage, the only way to investigate is by monitoring actual patients in acute care. This means that non-invasive but experimental technology is partially implemented

⁸ Views on discontinuing intravenous nutrition and hydration vary greatly across countries (Asai et al. 1999, Jennett 2002, Moratti 2010, Demertzi A et al. 2011, Beljaars et al. 2015).

in the care practice. In chapter 1, I have called these situations *liminal innovation practices*. Data is collected from real life patients who are simultaneously treated and, thus, subject to SFP whenever life-sustaining treatment is withdrawn. In this setting, the iteration of mistakes due to the lack of error signals happens in three different ways.

3.1. Established prognostic tests: including past mistakes in blinded research

First, we must acknowledge how the potential for hidden false positives in existing, established prognostic technologies influence the outcome of patients in contexts where novel tests are being evaluated.

To illustrate, consider novel research done to identify the prognostic relevance of continuous electroencephalogram (cEEG) with the prospect of including it in the multimodal neuroprognostic practice for patients after cardiac arrest. Assume that this research is done in hospitals where extra measures are taken to ensure that the treating physicians are blinded from the results of the study in every way possible so that practice and research remain relatively separate. For instance, my empirical observations (as described in chapter 4) showed that monitors can be turned off, facing the wall, covered with a sheet, or monitored on a different floor entirely. During this research phase, decisions on WLST are thus not based on cEEG but *other* prognostic indicators. Poor outcome can, for instance, be predicted after testing somatosensory evoked potentials (SSEP).

However, the false positive rate for absent SSEP indicating poor neurological outcome is estimated to be 7,7% (Amorim et al. 2018). This estimate is based on patient cases where life-sustaining treatment was *not* withdrawn and that were thus not subject to SFPs. The study adjusted for a WLST-rate of 80%. Although researchers therefore have good reason to believe those cases of false positives are there approximately 7,7% of the time, since they cannot identify them, they cannot remove them from their data set.

Inevitably, these false positives are incorporated in the research as if they were true positives. Therefore, even if no results from the preliminary findings on cEEG are considered in the decision-making because the treating physicians are blinded from them, the results of the research on cEEG is already subject to 7.7% of false positives that are cloaked by the decisions based on the previous prognostic test, namely, the SSEP test.

In essence, when the performance of a new prognostic tool is measured in reference to the current standard of care, its performance will always be constrained by the limitations of that standard. In this particular case, the standard is known to be flawed due to past instances of SFP.

3.2. Liminal innovation practices: including current mistakes

Unfortunately, the distinction between research and care practices in liminal innovation practices (Mertens 2018) does not generally hold as well as we would hope (Beauchamp and Shagai 2012). Since the patients are treated *while* research is being done and sometimes even by the treating physician, it is difficult to sustain the artificial barrier between research and clinical practice. In fact, it is likely, and often even required, that the treating physician *has* to include the tool under investigation in order to provide the best possible care to the patient. This is precisely where we must face the characteristics of liminal innovation practices.

Although blinding the treating physicians would help ensure the integrity of the research, this may not always be consistent with providing the best possible care to the patient. For example, even if efforts are made to blind the treating physician, like in the previous case, cEEG could also be used to detect epileptic activity which can subsequently be treated with anti-epileptic drugs. Once the treating physicians include the data for one treatment decision it is hard to leave that information out of the equation when making other treatment decisions. Moreover, as the research progresses, the pressure of having to make an uncertain decision without much to go on sometimes leads practitioners to explicitly seek out the data regardless of its research status, as it is considered 'better than nothing'. Thus, even if the physician in charge of prognosis is typically blinded from the administration of the novel test being tested, patient care may favor using some of the instruments involved in the novel test.

However, even if we leave these cases aside, systematically blinding treating physicians is not actually as common in these practices as one would hope. According to one review, only 9 out of 73 neuroprognostic studies reported blinding treating physicians from research results (Sandroni et al. 2014). In those settings, research-practice overlap is more likely to unfold as follows: The clinician in charge of prognosis and decision-making is not blinded to the instrumentation being tested. So, even if the established prognostic test is SSEP, an EEG machine that is part of a novel test may be visibly present in the room with the patient and the physician. This invites a natural form of confirmation bias (Worster & Carpenter 2008). Any hypothesis entertained about the evidential value of indicators in the new instrumentation affects not just the perceptions of the researchers but also the decision-making of the clinician (who may well be one of the researchers). So, an initial hypothesis about the deliverances of a new test affects treatment decisions, thus affecting the patient outcomes. In particular, if the hypothesis suggests a poor prognosis and this influences treatment decisions accordingly, the prognosis becomes self-fulfilling thereby spuriously confirming the hypothesis. So, even if the suspected positive was a false positive, this cannot be noticed. The evidence available shows the positive test result as a true positive. The consequence is that any deficit of specificity in the new test is made invisible. Quite appropriately, then,

we already see criticism of prognostication studies for lacking the sort of blinding that we have come to expect in other sorts of clinical trials (Sandroni et al 2014).

3.3. Continuous innovation: including new cases in ongoing research

Finally, a new problem arises as we move forward to improve the predictive capacities of new prognostic tools once they are established. Once a new tool like cEEG is integrated as part of the standard care practice, we are faced with a new challenge that comprises a combination of the previous two. On the one hand, the use of cEEG will inevitably affect the data of new prognostic studies. On the other hand, a self-fulfilling prophecy loop is inevitable if we continue to do research to improve cEEG's prognostic capacity.

To make this tangible, let us examine the novel use of cEEG for neuroprognostication after cardiac arrest more closely. Research shows that cEEG is a reliable indicator of good as well as poor outcome for approximately half of postanoxic coma patients when used in combination with SSEP (Ruijter, 2019). In several countries, cEEG is therefore accepted as a reliable addition to multimodal neuroprognostication of postanoxic coma (Horn et al. 2019, Nolan et al. 2021). As we discussed above, even if the study on cEEG was blinded, the predictions of cEEG have already been affected by the unacknowledged false positive rates of other prognostic tools that were used for decision making while the new tool was researched.

Once cEEG itself is used for end-of-life decisions, the outcomes of those patients are affected by potential false positives undetected due to the SFP, specifically caused by the prediction based on cEEG. Although employment of these new predictions again harms individual patients, this would not necessarily lead to another cycle of iterated mistakes⁹, were it not that research on cEEG's predictive value continues.

With the addition of machine learning techniques, the outcome of increasingly more patients can be predicted, up to 58% based on cEEG alone (Van Putten et al. 2018). Unfortunately, if new models are trained on the basis of data in which cEEG was included in the prognostic process, this form of bias does not only become more robust, it can further amplify. This way of using output as input is paradigmatic for feedback loops. As such, old mistakes lead to new SFPs which confirm the mistaken choices as correct and then go on to reinforce the prognostic method that produced the predictions. The result is an automated feedback loop of self-fulfilling poor outcomes for patients after cardiac arrest.

⁹ Once a new tool is investigated while cEEG is part of standard practice we are back to iterating the same mistake as in 3.1.

This feedback loop can easily be avoided by not including new cases, but how can the algorithm then be improved? Practitioners and researchers are again faced with a choice between either not basing clinical decisions for patients on the novel predictive tool (potentially the best prognostic tool we have) *or* not including new patients in research. Yet, by only including patient' deaths of whom life support has not been withdrawn, they would not be including very many patients at all.

In the three ways described above, undiscovered false positives are seen as true positives. This impedes improvement of new prognostic tests that inform the predictions of outcomes of future patients. The initial erroneous prediction does not remain simply unidentified after having caused harm to a single patient. It furthermore becomes a building block in the knowledge base from which the outcomes of more patients are predicted in the future. The result, in each case, is that false positive rates of novel tests will likely appear lower than they actually are. The earlier and the more often life-sustaining treatment is withdrawn, the worse this problem is. In section 5, I give guidance on how to improve quality assurance of predictions.

4. FACING NORMATIVE UNCERTAINTY. ARE WE LEARNING WHEN WE CAN?

While the lack of error signals or appropriate positive feedback is paramount to the epistemic issues SFPs pose, the issue is not always that error signals are not at all produced. Sometimes, the error signal - produced after an error is made - is simply not picked up. Error signals are not necessarily self-evident, nor do they always happen automatically. However, when SFPs produce error signals, the latter will be hidden from sight. By appearing true, these self-fulfilling predictions will again not call for accuracy checks.

Let us recall the clinical dilemma. Patient Yann whose outcome was predicted to be poor, died after WLST. If the prognostic test resulted in a false positive, we can't find out because no error signal is produced. We've discussed the epistemic and normative implications above. Patient Xander's treatment, however, was continued after his outcome was predicted to be good. By continuing treatment we're bound to find out Xander's eventual outcome and if the prognosis was mistaken, we should be able to notice the error signals.

In what follows, we will lay out how false negatives and positives can go undetected even after the patients survive and, as such, lead to self-fulfilling prophecies that, ultimately, inhibit learning. The following kinds of SFPs in neuroprognostication have thus far been neglected and require significant attention.

4.1. Normative evaluations of outcomes in neuroprognostication

The medical literature regarding SFPs has been focussed on the accuracy of the prediction; whether a patient predicted to have a good outcome, indeed has a good outcome and whether a patient predicted to have a poor outcome, actually has a poor outcome. However, the qualifications of ‘good’ and ‘poor’ are, by definition, normative evaluations. Yet, these normative evaluations refer to much more objective descriptions defined in the Glasgow-Pittsburgh cerebral performance categories (CPC) and Glasgow Outcome Scale (GOS)—see Table 7.3 below.

CPC	CPC description	GOS description (By Jennet and Bond, 1975)	GOS
1	Good cerebral performance / Good recovery		5
	Conscious, alert, able to work and lead a normal life. May have minor psychological or neurologic deficits	Resumption of normal life even though there may be minor neurological and psychological deficits	
2	Moderate (cerebral) disability		4
	Conscious. Sufficient cerebral function for part-time work in sheltered environment or independent activities of daily life	(Disabled but independent) Independent for daily living and able to work in sheltered environment	
3	Severe (cerebral) disability		3
	Conscious. Dependent on others for daily support (in an institution or at home with exceptional family effort)	(Disabled and dependent) Dependent for daily living in an institution or at home with exceptional family effort	
4	Comatose or (persistent) vegetative state		2
	Unconscious/unaware of surroundings, no cognition. No verbal or psychological interaction with environment	Patients who remain unresponsive and speechless for weeks or months until death	
5	Death		1
	Death by traditional cardiopulmonary criteria or certified to be brain dead	Agreement must be reached on what conditions should be met before ascribing death to brain damage	

Table 7.3 - Comparison of Glasgow-Pittsburgh cerebral performance categories (CPC) and Glasgow Outcome Scale (GOS) scores

To make sure they correctly assess the eventual outcome of surviving patients, researchers call the home of the patient (or the facility where the patient resides) and ask either the patient but often a family member or caretaker a series of questions. The answers allow an assessment that points to one of the categories. This assessment is often done 3, 6, and 12 months after discharge as patients tend to no longer improve past that point.

By dichotomizing these five categories of neurological outcomes in 'good' and 'poor', prognostication never actually predicts 'good recovery' or 'moderate disability'. The two are lumped together under the single denominator: 'good outcome'. Similarly, there is no prediction of 'death' or 'vegetative state' or 'severe disability'. The prediction is one of 'poor outcome' and it can be one of the three. That is a wide range of outcomes under a single umbrella. Especially since a predicted outcome under this umbrella has significant consequences. This dichotomy poses several problems, some of which I've discussed in previous chapters. However, a comprehensive discussion requires careful analysis and including all of them in this chapter will take us too far afield. Here, I aim to highlight a single problematic aspect: while the normative evaluation of these categories is not stable, it nevertheless impacts the way normative predictions can self-fulfill.

In chapter 3, we saw that these normative judgments are subject to change. They can vary from person to person and they are context-dependent. This is especially true for the category of severe disability which can include physical and/or cognitive disability; but even disorders of consciousness can be evaluated differently by, for instance, a religious community. Furthermore, the same person may evaluate the same outcome differently at different times in their lives, depending on their context, age, role, responsibilities, etc. Surely, losing all mobility will be valued differently by the twenty year old athlete than by the sixty five year old who is already in a wheelchair before hospitalization.

These variable evaluations of specific outcomes complicate our ability to correctly assess the accuracy of our predictions. And again, if we cannot identify potential mistakes, we cannot learn from them.

4.2. Undetected false positives and negatives due to categorisation issues

Researchers may evaluate a patient's actual outcome by asking a family member to evaluate the patient's short term memory or ask whether the patient can dress independently, speak clearly, make themselves something to eat or drink, or whether they are sufficiently mobile to join a trip to the store,... The answers to those questions will lead to significantly different information than when the patient (or their family) is asked whether they are happy, how they've felt since they recovered from the cardiac arrest, how they would value their own quality of life,... While the two sets of answers may be connected, the first are likely to indicate levels of physiological recovery, and the latter indicate normative evaluations.

Now imagine a patient, Mark, predicted to have a poor outcome. Despite the prediction, life sustaining treatment is not withdrawn because his family refuses

to give credibility to the physician's prediction. Although Mark's life-sustaining treatment is continued, he ends up with severe neurological disability. According to the categorization scheme the physicians refer to, this is indeed a condition that counts as a 'poor' outcome. So, from the standpoint of the physician who made the prognosis based on a positive test result, that prediction was a true positive. However, suppose that, from Mark's standpoint, life, even with the fairly severe condition, is still very much worth living and he considers it a 'good' outcome that he didn't pass away after the arrest.

Physicians and researchers may have good reasons for resisting such subjective evaluations but we'll get to that later. If, for now, we are inclined to think that we should privilege the patient's own evaluation of what is worthwhile or not, then we should say that the original prediction of poor outcome for Mark was a false positive, categories be damned.

Similarly, suppose Nicolas is a patient who survives with moderate disability. This is considered in line with the prognosis which predicted a good outcome, even if he experiences changes in his character that alienates him from his loved ones who no longer recognize the man they knew. Nicolas may become terribly depressed about his losses and may wish he had died instead. If he considers this a poor outcome but the factsheet puts him in the good outcome end of the spectrum, the error signal is not recognized and Nicolas' outcome, although a false negative to him, will be confirmed as a true negative.

Such assessments can thus result in a second SFP for patients that survive past the point where the first one takes place. Of course, they don't (at least not at this stage) physically affect the patient, changing their actual outcome like the previous SFP. These are *interpretative* self-fulfilling prophecies, meaning that the outcome is interpreted to be what it was predicted to be.

The misinterpretation, here, results from consolidating two very different steps as though they were one. While the prediction is made in normative terms, researchers assess the level of functional recovery based on rather objective measures of physiological recovery. These objective measures are then translated back into value judgments of 'good' and 'poor' in order to assess whether the predictions were correct. This is currently a necessary step, only because predictions are not made at the level of functional recovery, only at the level of a dichotomized categorization divided in 'good' and 'poor'. However, the dichotomy is based on normative factors.

In other words, the prediction of poor or good outcome was self-fulfilling for such cases, not because it caused the patient's physical condition to be different than it otherwise would have been, but because the way researchers assigned value to that physical condition matches the outcome they predicted.

An error signal or appropriate positive feedback would only be possible if the *actual* level of functional outcome can be compared with the *predicted* level of functional outcome. Since the prediction is a value judgement, it can only really be compared with the actual value judgment of the outcome. When an error is made, it is made in between the objective assessment of functional recovery levels and the translation of that objective measure to be 'good' or 'poor'.¹⁰

These errors are usually missed since they do not produce an error signal to the predictor *and* the mistake hardly matters to the actual patient who is often even unaware of the misevaluation. Moreover, as long as the interpretation does not affect policy, the patient is not affected by the misevaluation. Unfortunately, these unidentified false positives and negatives nevertheless have consequences for future patients as they are informing (or simply confirming) our future evaluations of good and poor outcome. Again, the ethical issue of SFPs pertains to our ability to accurately learn from failures and successes.

This shows why also identifying interpretative SFPs is essential because they make it difficult to correctly identify error signals as well as positive feedback and, as such, they inhibit learning. At least they do so temporarily, until mistakes accumulate over time and, as we will show, combined can form an error signal that is much harder to ignore.

4.3. Self-fulfilling predictions of normative futility and utility

Considering the variety in people's experience and valuation of neurological outcome, the pattern of our collective social judgement, given sufficient error signals, is bound to evolve. As previously mentioned, a large literature review showed that prognostic studies conducted before 2006 generally categorized 'severe (cerebral) disability', or category 3 on both the CPC and GOS scale, as a 'good' outcome. After 2006, however, the majority of the studies grouped the very same outcome with the other 'poor' outcomes (Sandroni & Nolen 2015, Sandroni et al. 2018). The authors' hypothesis is that a shift in our valuation of being independent caused a re-evaluation of the categories (see tables 7.4 and 7.5).

The significance of this change is profound. Somewhere along the line, sufficient error signals caused a shift in categorisation. Whether these error signals were accurately recorded, or just a consequence of interpretation mistakes remains to be investigated. What we do know, however, is that recently, error signals in the

¹⁰ To be clear, I do not argue that the patient's subjective normative evaluation is necessarily a better evaluation than one based on more objective descriptions of functional recovery. My argument is simply that when checking the accuracy of the predictions made, one cannot compare apples with oranges. Either normative judgments are compared with normative judgments or physiological judgments are compared with physiological ones.

other direction have led to some changes in policy in the Netherlands and, as of January 2019, funding was increased for intensive neuro-rehabilitation for patients surviving with severe disability and even disorders of consciousness (CPC4). It is clear that the existing normative evaluations are not yet accurately captured by the categories used to define neurological outcome. Especially the valuation of category 3, indicating severe disability, is found to be too diversified. And yet, the way practitioners dichotomize good and poor is essential because their normative evaluations of which outcomes are good or bad, can themselves lead to self-fulfilling prophecies.

Before 2006:

CPC	Description	GOS
1	Good cerebral performance / Good recovery	5
2	Moderate (cerebral) disability	4
3	Severe (cerebral) disability	3
4	Comatose or (persistent) vegetative state	2
5	Death	1

Table 7.4 - Distinction of good and poor outcomes in Glasgow-Pittsburgh cerebral performance categories (CPC) and Glasgow Outcome Scale (GOS) scores before 2006

After 2006:

CPC	Description	GOS
1	Good cerebral performance / Good recovery	5
2	Moderate (cerebral) disability	4
3	Severe (cerebral) disability	3
4	Comatose or (persistent) vegetative state	2
5	Death	1

Table 7.5 - Distinction of good and poor outcomes in Glasgow-Pittsburgh cerebral performance categories (CPC) and Glasgow Outcome Scale (GOS) scores after 2006

In order to provide a careful analysis of the actual SFPs caused in different countries by these categorisation issues, further empirical investigation is required. In the meantime, we can give one example of how a prediction of a normative evaluation can cause that evaluation to take hold, actually affecting the outcomes of all patients in the patient group.

When value judgments of a neurological outcome impacts treatment decisions regarding WLST, it increases or reduces the surviving population who has such an outcome. For instance, a greater scope of the evaluation ‘poor outcome’ (incl. category 3) combined with a policy of WLST for this evaluation, immediately leads to a reduced population of living patients after postanoxic coma—*especially* a reduced population of living patients in category 3. Smaller or larger patient populations demand different care policies and these policies affect the outcomes of the patients in two ways. When discussing collective SFPs in the previous chapter, I

argued that the sources of self-fulfillment can be interpretative or substantive. Here too, patients may be perceived as worse or better off, and their actual quality of life can be reduced or increased, mirroring the original value attached to their outcome.

Imagine the following scenario. The prediction 'Patients with severe disability are patients with a poor outcome' (as in Table 7.5) is generally accepted and given credibility. Again, employment of the prediction by ways of WLST causes a first self-fulfilling prophecy for all those whose outcome was predicted to be poor, including those with a predicted severe disability. Due to this first SFP, fewer patients with severe disability survive. Imagine that, due to the decrease in population and the increase of 'perceived' futility, medical treatment and support in the form of neurorehabilitation, therapies, research, etc for this population is then reduced. Due to reduced resources the patients who survive with severe disability now have a poor (or worse) outcome.

As such, when value judgements are systematically leading to mistaken categorisations this can have a self-fulfilling effect on future patients and all members of the patients group. Further empirical investigation may be required to fully grasp how such SFPs precisely take place for patients with severe disability. However, clinical research in the Netherlands for patients with disorders of consciousness gives reason to believe something like this has already happened. In her study of the unresponsive wakefulness syndrome (UWS, often referred to as vegetative state (VS)) van Erp describes the following process as a 'vicious circle':

Figure 7.5: Vicious circle of unknown epidemiology, suboptimal care, and tendency to discontinue treatment (van Erp 2020)

Step 1: the group of patients seems small and recovery is rarely witnessed, especially by those responsible for care in the acute phase. Step 2: because of the supposed small numbers and modest chances of meaningful recovery, care is organized ad-hoc, resulting in misdiagnosis, shattered expertise, lack of specialized rehabilitation and family counselling. Step 3: decisions about whether or not to continue life-supporting treatment are made without solid diagnosis or scientifically sound prognostics, often by a physician without knowledge of the possible long-term outcomes and without a concrete roadmap to adequate post-acute and long-term care. This brings us back to step 1: the number of patients and their chances of recovery remain small due to a tendency to discontinue life-prolonging treatment, while those who survive continue to receive suboptimal care. (van Erp 2020, 92-93)

Fortunately, the accumulation of error signals in chronic care triggered a learning process that resulted in recommendations for how to break the vicious circle. In the same vein, the limited but meaningful collection of error signals in chronic care for people with severe disability offer two insights. First, that a patient with severe disability can have a 'good' outcome. Second, given immediate intensive neurorehabilitation, a patient's outcome might be significantly better. Such insights are crucial when, as we saw in chapter 6, a reversal of the initial prediction can again have a self-fulfilling effect.

Imagine the following scenario. The prediction 'Patients with severe disability are patients with a good outcome' (as in Table 7.4) is generally accepted and given credibility. Again, employment of the prediction by ways of WLST causes a first self-fulfilling prophecy for all those whose outcome was predicted to be poor, no longer including those with a predicted severe disability. Due to this first SFP, more patients with severe disability survive. Imagine that, due to the increase in population and the decrease of 'perceived' futility, medical treatment and support in the form of neurorehabilitation, therapies, research, etc for this population is then increased. Due to increased resources the patients who survive with severe disability now have a good (or better) outcome.

These scenarios are too categorical to be completely realistic and what really happens is likely more diverse and complex than how I describe it here. There are, furthermore, other ways in which such predicted judgements can become self-fulfilling. For instance, due to lower or higher survival rates, the general public may become more alienated or more familiar with such patients, treating them differently in ways that affect their actual wellbeing or quality of life. As mentioned in the previous chapter, similar arguments have been made for autism, dwarfism, and even deafness. In any case, accurate outcome assessments probably lie somewhat in between the two predictions, with some patients surviving with severe disability having good outcomes and some surviving with severe disability having poor

outcomes. Responsible prognostic innovation requires correctly identifying which patients are which.

5. GUIDELINES: CAN WE LEARN FROM HIDDEN MISTAKES IN THE FACE OF PHYSIOLOGICAL AND NORMATIVE UNCERTAINTY?

Generally, SFPs inhibit learning because their self-fulfillment makes reliable feedback difficult. For neuroprognostication in acute care, it is impossible to get an error signal once the prediction self-fulfilled due to WLST¹¹. However, error signals regarding physiological futurity predictions are still produced in cases where life sustaining care wasn't withdrawn after a prediction of poor outcome was made and the patient ended up with a good outcome. Naturally, these cases are rare in countries where WLST is common. Precisely for that reason, they deserve extensive scrutiny. Similarly, patients whose outcomes were predicted to be good but ended up poorly, need to be examined. These cases may teach researchers precisely what they are otherwise missing.

Recently, based on these exceptional cases of detected false positives, efforts have indeed been made to estimate the prevalence of missed error signals. For instance, an estimation of the *prevalence* of false positive rates for SSEP was made by Amorim and colleagues (2018). Similar studies should be done for other prognostic indicators, even if these estimates of frequency do not tell us anything about the *nature* of the false positives.

Responsible innovation requires a more careful approach with respect to each of the three ways in which SFPs distort research in acute care. The following three guidelines are relevant not only to researchers, but also to peer reviewers who are advised to integrate this guidance into their quality assurance process:

1. To avoid undiscovered false positives from established technology propagating to new technology, it is imperative to take into account the old test's unbiased estimates of false positives, adjusted for WLST (if only statistically determined), when assessing the new test.
2. When possible, treating medical staff should be completely blinded from neuroprognostic studies. Whenever research data are explicitly requested to inform treatment of a patient, responsible research practice requires taking that patient out of the data set.
3. Ideally, new models should not be trained on data influenced by the use of current tests. In research where models have been trained on tainted data, this problem should be clearly flagged and strategies for mitigation should be discussed.

¹¹ Some postmortem studies aim to improve quality assurance checking after the death of the patient. (van Putten et al. 2018, Endisch et al. 2020)

In cases where adherence to these guidelines poses a challenge to collecting sufficient research data, alternatives ought then to be explored. We recommend considering collaborations with researchers in other countries where there are different practices regarding WLST, keeping in mind other limitations these contexts may have. For instance, countries where WLST is uncommon typically struggle with stronger resource limitations in intensive care due to the higher demand for life supporting technologies. When studying cases abroad, it is important to evaluate those that compare to the same standard of care as those of the home country.

With regard to missed errors in chronic care, much room for improvement currently lies in optimizing coordination between acute and chronic care. This will help identify the kind of error signals with regard to predictions of normative futility that are detectable. Furthermore, assessment of actual physiological outcomes should not be compared with predicted normative assessments of outcome. Again, the key here is to not compare apples with oranges. Either normative judgments are compared with normative judgments or physiological judgments are compared with physiological ones.

While physicians may have good reasons to stick to rather objective measures of functional recovery levels, these simply do not allow proper assessment of the prediction's accuracy. Predictions must either reflect these functional recovery levels, or, if neurological outcomes continue to be predicted at the normative level of 'good' and 'poor', subjective value judgments may be the better way to assess actual outcome. If the latter is the case, collaborative, coordinated care will be required to make sure the evaluations of patient outcomes are appropriately sensitive to error signals as well as to adequate positive feedback. Further empirical study is furthermore required to study the different ways in which normative predictions self-fulfill in different contexts. Meanwhile, it would be wise to recognize outcomes that fall under the most ambiguous normative evaluations as unpredictable or, at least, unknown for the time being. If we can admit that there are patients for whom no prognostic information is available and for whom the physiological outcome remains uncertain, surely we can do the same for the normative evaluations for which our predictive capacities are lacking.

6. CONCLUSION

There are two self-fulfilling prophecy mechanisms at work in neuroprognostication. The first SFP of physiological futility is widely recognized but poorly understood. This SFP makes detection of error signals really difficult, if not impossible.

Furthermore, when such SFPs are at play, positive feedback cannot be considered an appropriate source for learning. A second, normative, SFP is not recognized and although it produces error signals, these are often missed. On the one hand, attempts to detect error signals are not always made. On the other hand, when efforts are made, the procedures in place are flawed.

Due to both the inability to detect or correctly identify false positives and negatives, as well as the inability to rely on positive feedback as indicative of true positives and negatives, both SFPs inhibit learning and diminish the ability to accurately predict outcome for future patients. Since evaluation of past, current, and future patients is so intertwined, prognostic innovation practices risk being subject to self-fulfilling prophecy loops.

In general, it is worth reemphasizing the upshot of the core challenge posed by SFPs: When predictions bring about their own fulfillment, researchers can never rely upon, as genuinely informative, any feedback they receive about the accuracy of those predictions.

Learning is, however, a fundamental part of research and innovation. If we have an innovation process, or research method, that limits how much we can learn from our successes and failures, that research fails to do what research is supposed to do, namely produce accurate learning. Responsible learning requires attention not just to the mistakes that are most apparent to us. It also demands that we heed the mistakes that are hidden from us, as with prognoses that are self-fulfilling.

Conclusion

pitfalls and opportunities for responsible neuroprognostic innovation

Neuroprognostic innovation is driven by the lack of tools required to answer the question whether continuing life-sustaining treatment is the best way to care for a patient in coma who has likely suffered substantial brain damage. Essentially, decision-makers want to avoid discontinuation of treatment for patients who have good chances of recovery while simultaneously avoid continuation of life-sustaining treatment for patients who have no substantial chance of recovery. Failure to accurately predict the eventual outcome leads to two risks of harm: a) deaths of patients who could have otherwise recovered with a good quality of life on the one hand, and b) patients who survive in conditions they may experience as worse than death, on the other. Decisions to withdraw treatment entails one risk, decisions to continue treatment risks the other. This is the clinical dilemma that decision-makers are faced with. Since existing neuroprognostic practices are so precarious, the search for better prognostic tools is commendable. With this research, I give guidance on *how to ensure responsible innovation of prognostication in postanoxic coma*¹.

RESPONSIBLE INNOVATION OF PROGNOSTICATION AFTER CARDIAC ARREST

Investigating *how cEEG-enabled prognosis may affect the current concerns in care for postanoxic patients in coma within the Netherlands*, first required identifying the current concerns, to what extent they are currently addressed, and how? Due to uncertainty, the realization of even the most basic values in medicine such as pursuing the patient's health and wellbeing and respecting their autonomy are currently jeopardized. Decision-makers want to avoid harm but the uncertain options they have may cause harm anyhow.

As a way to address the existing concerns, innovative prognostic tools, like cEEG, aim to reduce the pervasive uncertainty in order to facilitate the decision-making process and reduce the risk of harm. However, the empirical fieldwork presented in chapter 2 revealed that the uncertainty that decision-makers struggle with is not simply about not knowing what eventual physiological outcome to expect. Often, there is uncertainty about the way that outcome will be experienced by the patient and whether or not this would be acceptable. In chapter 3, I defined these two types of uncertainty as *physiological and normative uncertainty*. As such, they mirror the two types of medical futility judgments they are complicating. Uncertainty

¹ Research questions are marked in bold.

about physiological futility is *uncertainty whether the proposed treatment can physiologically achieve the (foreseen) desired effect*. Uncertainty about normative futility, however, is *uncertainty whether the (foreseen) physiological effects of the treatment constitute an acceptable outcome*.

Since neuroprognostic tools are focused only on reducing physiological uncertainty, I further investigated the nature of normative uncertainty, the concerns that emerge from it and how it could be addressed in endeavors of responsible neuroprognostication. Currently, neuroprognostic tools have fixed normative evaluations built in since the different outcome categories are dichotomized into 'good' and 'poor'. This classification asserts a normative certainty and stability that is at odds with the normative uncertainty experienced in practice. To address this uncertainty, clinicians in the practices I observed typically try to obtain information on the patient's views regarding their quality of life and what they would consider acceptable. However, this way of addressing normative uncertainty again assumes a stability in individual evaluation of certain neurological outcomes and while this may be true for some outcomes, it is not the case for more ambiguous outcome categories like severe neurological disability. As described in chapter 3, normative uncertainty, especially for these cases, is not merely due to a diversity in views on quality of life. Even individual views are very dynamic. Whether specific forms of neurological disability are considered acceptable depends on factors intrinsic to the patient as well as on external factors. External factors like the availability of appropriate care (or lack thereof), both in the personal sphere as well as on an institutional level, can furthermore strengthen (or weaken) a person's internal orientations like their resilience and optimism. These variables make it difficult to predict whether some outcomes can be considered good or poor and thus they complicate the reduction of normative uncertainty. By maintaining the existing normative assumptions, innovative tools like cEEG, keep the problems caused by normative uncertainty under wraps.

While limited in addressing normative uncertainty, new prognostic tools may still realize the reduction of physiological uncertainty. And, if prognostic information is less uncertain, it might provide more clear-cut information for medical professionals, offer better guidance for decision-making, and be easier to communicate to relatives. These three improvements can make it easier to decide whether to continue or withdraw life-sustaining treatment, at least for some patients.

Further analysis of practices in which cEEG was implemented for research purposes shows, however, that the greater prognostic clarity that cEEG offers does not automatically lead to better understanding of the patient's eventual outcome, better communication with the family, or better clinical decision-making. ***To develop the cEEG-based prognostic technology in such a way that it addresses existing***

concerns and takes into account potentially undesirable effects, extensive attention must be paid to the way cEEG will be implemented in practice. As shown in chapter 4, the extent to which cEEG can address existing concerns or cause undesirable effects will therefore, to an extent, be context-dependent.

Generally speaking, practices that include cEEG must be careful about how the patterns that cEEG produces are interpreted. This will require deliberate choices about who is allowed to interpret and how. Opting for visual interpretation will require proper training of experts. When opting for computer-assisted interpretation, transparency towards the experts about the way patterns are interpreted will be crucial. Especially when the results of two very different computer analyses are represented in the same way, this can be misleading. Furthermore, some forms of machine learning will make transparency with regard to interpretation difficult to achieve (e.g. when a neural network is trained to independently predict without being fed specific features).

The influence of the user-interface on communication should not be underestimated either. The visibility of the screens and the design of the interface should be such that they prevent non-expert clinical staff and family members from jumping to conclusions prematurely. The use of green and red colors especially (and traffic light models more generally) tend to inform people at times and in ways that are undesirable. Such color indications also have prescriptive power, implying whether to continue or stop. These can be deliberately intended to inform decisions to continue or discontinue treatment, or merely intended to be transparent and informative with unintended effects on the subconscious. Either way, these prescriptive colors may raise (false) hope or cause (unwarranted) alarm, complicating communication and decision-making later on. The interface can also complicate communication in cases where no prediction can be made. An unintended consequence of cEEG's ability to predict both good and poor outcomes is that, depending on the interface, a third, undetermined category emerges. Making explicit the category of patients for which no prediction is available can lead to negotiation processes. While this may be desirable in some cases, it can also be counterproductive and undermine the relationship between the clinicians and the patient's next of kin.

Chapter 4 lays out in detail what undesirable effects specific to cEEG need to be considered depending on the context in which it will be implemented. If these are taken into account, the addition of cEEG will significantly strengthen the predictive value of the multimodal prognostic process. Especially its abilities to identify good outcome and to increase identification of poor outcomes in aggregate with existing tests, can strengthen the tests overall prognostic value. Furthermore, cEEG's performative effects show promising potential for further prognostic and

even therapeutic² innovation. By bringing to light patterns that are indicative of remaining uncertainty, cEEG helps identify patients for which the outcome is neither clearly good, nor clearly poor. More research can be done to better address the physiological and normative uncertainty for patients with these outcomes.

Some of the ways in which cEEG impacts current concerns, or fails to do so, are not specific to cEEG but are reflective of problems inherent to the way neuroprognostic innovation in general is constructed. I have already pointed to the concerns regarding normative uncertainty that are pervasive, regardless of the innovative prognostic tool under investigation. Another such pervasive problem in neuroprognostic research is that of self-fulfilling prophecies. I will discuss below how both self-fulfilling prophecies and the entanglement of the two types of uncertainty jeopardize responsible neuroprognostication innovation. Before tackling these two general problems and their implications, however, I would like to review my findings from an even wider perspective.

RELEVANCE FOR CLINICAL INNOVATION AND BEYOND

In the introduction, I promised to distill *what findings can be generalized to neuroprognostic innovation practices for patients in coma and, broader still, to innovation practices in clinical settings*. I will start with the latter, discussing the broader relevance of my work and then go back to the more general conclusions relevant to neuroprognostic innovation specifically.

Based on my observations of cEEG as a case study of clinical innovation, I concluded that the existing assumptions in the field of responsible innovation (RI) were not consistent with those of the clinical innovation practices I observed. In response, I developed the notion of *liminal innovation practices*, in chapter 1, which I described as *innovation practices defined by continuous transition on both sides of the threshold between experiment and implementation, between research and practice, and between past and future innovations*. Such innovation practices are so significantly distinct that they require a special kind of assessment. First, innovations at the center of such practices typically have different characteristics than the 'emerging technologies' often assessed in RI. Second, the innovation trajectory is much more malleable, almost contradictory to what is usually perceived to be the case. Since adjustments and even retractions can be made long after the technology has been implemented, there is no so-called 'Collingridge dilemma'. Due to their specific characteristics, I concluded that these innovation practices are best assessed

² The use of cEEG helped identify epileptic activity, for instance, which allowed the testing of anti-epileptic drugs. The value of other innovative therapeutic interventions such as early brain stimulation can likely also be assessed through cEEG monitoring.

by reorienting RI approaches away from anticipation of the unknown and uncertain, and returning to observation and assessment of the known and predictable. Towards the end of chapter 1, I suggested that since 'other innovation practices in clinical settings have similar features that contradict the common assumptions in RI, it is likely that many of them are liminal, too' but I had no empirical evidence to back this claim. I do now. During the global pandemic, makeshift test-centers and the way COVID-19 vaccines were developed proved to be (and are still proving to be) crown examples of liminal innovation practices in medicine across the globe.

These crisis-induced examples extend far beyond medicine too. Innovating the way digital test results are communicated and standardized for international verification and even the way emergency legislation is quickly implemented, retracted, adjusted, reimplemented, etc. exemplify the kind of liminal innovation practices for which RI-assessment is indispensable. A recent publication, analysing the developments of the pandemic on work-practices, argues that crises 'disrupt established practices, existing ways of doing things are interrupted and called into question' in ways that 'produces opening for liminal innovation, a process entailing iterative experimentation and implementation' (Orlikowski & Scott, 2021). In their article, Orlikowski and Scott highlighted the characteristics of liminal innovation that they found to be especially instructive to their work:

Our emphasis on liminal practices builds on the ideas of Mertens (2018), who developed the notion of liminal innovation as a set of practices that entail 'a seemingly never-ending transitional phase where one innovation serves as a building block, or a kick starter, for another one' (p. 290). That is, liminal innovations flow repeatedly 'between experimentation and implementation' (p. 280), being enacted in ongoing practices where 'partial implementation is a condition for experimentation' (p. 288). In this view, liminal innovation practices are open-ended, fluid, and flexible, ensuring that 'innovations remain malleable, even after implementation' (p. 286) (Orlikowski & Scott 2021, 2-3).

Although I said, in the original publication, that I could not be sure 'whether liminal innovation practices exist outside of clinical settings', I can now say with confidence that they do. Since its publication in 2018, several scholars have picked up on the concept in their publications and I have been approached by others who are now using it for their work. The diversity of cases for which liminal innovation practices appear to be relevant is remarkable. One scholar is currently working on a waste water recycling project in the United States. They thought that 'the technology [they are] working on assessing—the recycling of wastewater into drinking water—could best be described as liminal.' This led to [them] seeing if anyone had described some forms of innovation as liminal. After finding the 2018

article they thought that 'the issues [I] raise and approach [I] take will be useful for engaging with recycled water promotion projects in the desert southwest' (personal communication, 2020). While I could not foresee the uptake, I did assume that 'other scholars [would] recognize these liminal characteristics in their own RI research'. I am thrilled to confirm, two and a half years later, that this is indeed the case.

Similarly to how I felt about liminal innovation practices at the time, I now feel about the insight my colleagues and I have published on self-fulfilling prophecies (SFPs). While I will discuss my analysis of SFPs pertaining specifically to neuro-prognostic innovation below, I would like to emphasize the relevance of the descriptive framework of SFPs for medicine in general, presented in chapter 6. While I have only scratched the surface of the epistemic and normative implications for medical practice and research, I have laid the groundwork for myself and others to analyse different kinds of SFP in this field. This achievement was made possible also thanks to my collaboration with Owen C. King with whom I developed a significant part of the work on SFPs, more broadly construed, in practical and automated prediction. The latter, I believe, will be of great value for policy makers and the artificial intelligence (AI) research community. Especially considering the rampant rise of AI-use in predictive analytics and the increasingly pervasive and socially significant technologies in which these are embedded, artificial cognition processes will require extensive ethical and epistemic analysis. The work on SFPs shows that proper evaluations of practical and automated prediction, whether in medicine or elsewhere, will require quality assurance that looks beyond accuracy, as self-fulfilling prophecies are always, necessarily, true. In medical research and practice, this will be especially relevant for prognosis and, increasingly, for early diagnostics and predictive genetic testing, especially now that there has been renewed focus on personalized medicine and early detection and prevention strategies. Furthermore, because of the difficulty to uphold the research-practice distinction in clinical research, SFPs in medicine have serious epistemic implications that will complicate our understanding and ability to learn. I further expand on these implications in a moment.

Finally, since prediction is almost always done in response to a degree of uncertainty, distinguishing between different kinds of uncertainty that require different kinds of prediction is essential. Considering the impact of transformative SFPs, and reflexive predictions more generally affecting outcomes, we should be cautious about predicting disease and risk of disease. In order for a (prospective) patient to decide whether or not to take action in light of such predictions, the patient must also know more about the way they will experience specific outcomes. Furthermore, 'knowing of risk' may itself have positive or adverse effects on our health that may or may not outweigh the benefit of remaining with uncertainty.

Since knowing about a potential development will not help us in knowing how we will manage or experience that development, I suspect that distinguishing between normative and physiological uncertainty, and addressing both, will become essential in medical prognostics, early diagnostics, and predictive genetic testing.

These general insights require more investigation and, as such, they provide fertile ground for further, exciting research. For now, I am confident to present these broader insights as being deduced from and grounded in the more specific conclusions I drew from my analysis of responsible innovation in neuroprognostication. I will now expand on the general epistemic challenges that neuroprognostic innovation practices, regardless of the tool under investigation, pose for patients in coma.

EPISTEMIC PITFALLS FOR NEUROPROGNOSTIC INNOVATION

Generally speaking, when researchers seek to innovate an existing practice or technology, they aim to improve what is already there. Improving means both learning from mistakes and learning from successes. Learning as such is essential to innovation. Precisely for this reason, I have taken a closer look at the data collection and knowledge production that lie at the core of the development of a new prognostic tool. Following my suggestions in chapter 1, I conducted multicontextual exploratory research. Looking at different practices revealed the variety in epistemic processes involved in prognostication, how they differ and where they overlap. Additionally, it further confirmed the importance of evaluating epistemic processes in innovation.

DIFFERENT UNCERTAINTIES DEMAND DIFFERENT PREDICTIONS

My investigation revealed several epistemic pitfalls that indicate different ways in which the learning processes are flawed. The first main epistemic problem in postanoxic neuroprognostic innovation is that the uncertainty that decision-makers face and subsequently try to address has two components that are often confounded: physiological and normative uncertainty. When I discussed the existing concerns above, I already showed that, in practice, two different questions need to be answered: 1) what kind of neurological outcome can be expected for this particular patient? 2) will that outcome offer the patient an acceptable quality of life? Failure to predict either the eventual neurological outcome (due to physiological uncertainty) or the way that outcome will be experienced by the patient (due to normative uncertainty), entails risk of harm; either unnecessary death or unacceptable survival.

As said, prognostic technologies can, at best, only answer the first question and, thus, offer more certainty only regarding the physiological component. This limitation in itself need not be problematic. If this were the only problem, better neuroprognostic tools would nevertheless offer great improvement. After all, an answer to one of the questions is better than no answers at all. However, the way that neuroprognostic tools are researched and developed, has normative assumptions built in. First, the answer to the second question—whether a specific outcome will offer the patient an acceptable quality of life—is already assumed. The way that different outcomes have been dichotomized into ‘good’ and ‘poor’ outcomes implies that there is no normative uncertainty to speak of. Second, while a prognosis considered ‘poor’ can still be ‘trumped’ by relatives arguing for a different view of what is acceptable or not, the embedded but hidden categorization makes it difficult to know what precisely to argue against. Third, as I explained above (and in chapter 3), these individual normative judgments are also quite unstable. The preferences of Dutch citizens with regard to predictions of both physiological as well as normative futility judgments, presented in chapter 5, further indicate that the assumed normative answers are deeply contentious. All this makes holding on to fixed dichotomies especially troublesome.

Responsible innovation of neuroprognostic practices demand at the very least more transparency about the inscribed normative evaluations of outcomes so that decision-makers can better understand and communicate what a ‘good’ or ‘poor’ prognosis may entail. To this end, better coordination between acute and chronic care will be needed to get a more accurate picture of what can be expected. To ensure responsible innovation, however, neuroprognostic practices require pulling apart the two kinds of uncertainty. Research on new neuroprognostic tools must focus on answering the physiological question and the physiological question alone. In practice, this means that studies should not aim for prognostic results that predict ‘good’ or ‘poor’ quality of life but for results that describe the actual physiological outcomes. This can be done by referring merely to the cerebral performance categories (CPC) or corresponding categories of the Glasgow Outcome Scale (GOS). If more fine-tuned technology can be developed, one could refer to more refined categories like those of the extended CPC (CPCe). Acknowledging the limitations of *any* description, even these descriptions will never be completely value-neutral. Yet, they do not entail such strong normative evaluation as ‘good’ or ‘poor’.

In this way, at least, the physiological and normative issues and the respective challenges that they pose can be addressed individually and get the appropriate attention they deserve. This will be especially important for patients whose neurological outcome is categorized as severe neurological damage, otherwise known as CPC3. This category contains the most diverse outcomes which are

subject to the most diverse normative evaluations. The recent changes in CPC3's evaluation in prognostic research as well as the results of our survey in chapter 5 confirm this. The category currently represents the diversified threshold between outcomes that are widely acknowledged as being clearly acceptable and those widely acknowledged as being clearly unacceptable. A first step forward would be to recognize that we may simply lack the relevant predictive information to know how outcomes that fall under CPC3 will be valued.

NORMATIVE AND PHYSIOLOGICAL SELF-FULFILLING PROPHECIES

The other main epistemic problem that (although partially acknowledged by many researcher-practitioners) remains unaddressed in many neuroprognostic innovation practices is that of self-fulfilling prophecies. I expand on this problem in chapter 7 and show that self-fulfilling prophecies (SFPs) are also at work on the physiological as well as the normative level. Pulling the two uncertainties apart, as advised previously, will also help better address the challenges that SFPs pose.

As stated previously, in deciding how best to proceed with the care for a particular patient, decision-makers face two kinds of uncertainty. Hence, neuroprognostication requires two predictions, rather than one. One uncertainty requires us to predict the physiological futility or utility of continuation of life-sustaining treatment while the other uncertainty requires a prediction regarding the normative futility or utility of continued treatment. While innovation is mostly focused on the first prediction, reliable neuroprognostication requires that our predictive capacities regarding both predictions improve. Unfortunately, self-fulfilling prophecies in both respects hinder our capacity to learn from mistakes and successes.

When a prediction is made about what kind of physiological outcome can be achieved when life-sustaining treatment is continued, making the prediction and acting upon it can fulfill the prediction. When a prediction of poor neurological outcome leads to withdrawal of life-sustaining treatment and the patient dies as a result, the prediction is realized due to the prediction and the way it was employed. When predictions entail their own truth, researchers aren't in a position to check whether the grounds for their prediction were right at the time. This pollutes and distorts the data collected and research observations that inform the next steps in the innovation process. Therefore, as long as we assess predictions based on their predictive accuracy, catching errors will be difficult.

Similarly, some normative predictions about whether a specific outcome will offer the patient an acceptable quality of life, may have a self-fulfilling effect. In medical research and practice, variation in normative evaluations of prolonged

disorders of consciousness that go against the research standard, are often interpreted in practice as cultural, religious, or a result of disbelief that the outcome is irreversible. Variations regarding severe disability (CPC3), on the other hand, while acknowledged as reasonable by *any* standard, disappear in practice due to the research standards that include this category as a poor outcome. In practice, this leaves little room to renegotiate the normative valuation of a specific outcome, especially when people may already face a lot of uncertainty about their physiological evaluations. Furthermore, the variations in evaluations of expected quality of life are often understood to be fixed and merely subjective due to value diversity. However, changes in the patients' circumstances, their flexibility, and resilience, and that of their immediate environment, as well as institutional differences can impact the way they ultimately experience and evaluate their quality of life. Thus, even leaving normative evaluation of outcome to subjective assessment without further reference to other factors that could impact the patient's ultimate valuation seems inadequate. This confirms that expected quality of life cannot merely be tied to physiological outcome alone and I suggest that there is room for contextual and institutional improvement that can change the normative evaluations of patient outcomes. I furthermore argued, in chapter 7, how assumptions regarding what physiological outcomes will result in acceptable or unacceptable quality of life, can impact these contextual factors in self-fulfilling ways that make the outcome increasingly match the quality of life predicted. Here, again, errors are masked by the predictions' truth. But because the truth is mostly due to interpretative fulfilment, checking the actual evaluation could reveal errors and give opportunity for correction. Crucial, here, will be to compare physiological evaluations with physiological ones, and normative evaluations with normative ones. While this may seem self-evident, the existing confounding of the two types of evaluations complicates such comparison.

RESPONSIBLY RESPONDING TO EPISTEMIC PITFALLS

Clearly, my epistemic critique of neuroprognostic innovation does not pertain only to a single moment in the evolution of prognostic technology. As seen in chapter 4, whether the implementation of cEEG, specifically, can be considered responsible depends to a large extent on how exactly it will be designed and used. Yet, as the newest addition to multimodal prediction after cardiac arrest, cEEG is no more or less inherently flawed than any other technology that built on the shoulders of its predecessors and that is thus based on epistemic assumptions that have not been adequately scrutinized.

While a new prognostic technology like cEEG offers a promising step in the right direction, I have shown that there are several ways in which the learning processes

on which neuroprognostic innovation is based, can be flawed. First, many old assumptions are transferred to any new prognostic tool that is being developed. Second, these transfers may remain undiscovered and therefore unacknowledged. This inhibits responsiveness. For instance, innovation in neuroprognostication aims to reduce uncertainty. And yet, the development of a new tool does not prevent the adoption of existing errors in prediction of physiological futurity like missed false positives. Nor does it prevent the adoption of potentially mistaken normative predictions such as what neurological outcomes will be experienced as acceptable or unacceptable.

Looking at different practices that are based on different assumptions helped reveal the importance of evaluating epistemic processes in innovation. When a new tool is placed in a different context, different lessons will be learned. In the same vein, when a person changes contexts, they learn other lessons. Perhaps the most general lesson of this thesis, regarding how to responsibly innovate, is that learning processes in innovation require much more epistemological investigation than they commonly receive. My conclusion is that the following three steps are fundamental for responsible innovation in prognostic technology development: a) understanding how learning takes place and b) evaluating the knowledge production and learning processes as well as their failures in order to c) subsequently be responsive to these insights.

While self-fulfillment impedes proper assessment of newly gained understanding, correcting for epistemic errors due to data pollution or the impact of unfounded assumptions is especially difficult because neuroprognostic research and innovation take place in clinical settings. These make it impossible for clear research-practice distinctions to be upheld. Yet, as suggested in chapter 1, while the threshold between implementation and experimentation in these liminal innovation practices causes epistemological difficulties, such a threshold also leaves openings for analysis and amelioration of the epistemic challenges it gives rise to. Indeed, responsible innovation demands we do not miss these learning opportunities that allow responsiveness.

References

- Admiraal MM, Horn J, Hofmeijer J, Hoedemaekers CWE, van Kaam CR, Keijzer HM, van Putten MJAM, Schultz MJ, van Rootselaar AF (2020) EEG reactivity testing for prediction of good outcome in patients after cardiac arrest. *Neurology*, 95(6):e653-e661
- Akrich M (1992) The De-Description of Technical Objects. In: W. Bijker W and J. Law (eds) *Shaping Technology/Building Society: Studies in Sociotechnical Change*, 205–224. Cambridge MA: MIT Press.
- Amorim E, Ghassemi MM, Lee JW, Greer DM, Kaplan PW, Cole AJ, Cash SS, Bianchi MT, Westover MB (2018) Estimating the false positive rate of absent evoked potentials in cardiac arrest prognostication. *Critical Care Medicine* 46(12):e1212-e1221.
- Asai A, Maekawa M, Akiguchi I, Fukui T, Miura Y, Tanabe N, Fukuhara S (1999) Survey of Japanese physicians' attitudes towards the care of adult patients in persistent vegetative state. *Journal of Medical Ethics* 25(4):302-8.
- Bagheri A (2008) Regulating Medical Futility: Neither Excessive Patient's Autonomy Nor Physician's Paternalism. *European Journal of Health Law*, 15(1):45–53.
- Balouris SA, Raina KD, Rittenberger JC, Callaway, CW, Rogers JC, Holm, MB (2015) Development and validation of the Cerebral Performance Categories-Extended (CPC-E). *Resuscitation* 94:98–105.
- Barben D, Fisher E, Selin C, Guston DH (2008) Anticipatory Governance of Nanotechnology: Foresight, Engagement, and Integration. In E. J. Hackett, O. Amsterdamska and M. Lynch (eds) *The Handbook of Science and Technology Studies*, Third Edition, 979–1000. Cambridge, MA: MIT Press.
- Beauchamp TL and Saghai Y (2012) The Historical Foundations of the Research-practice Distinction in Bioethics. *Theoretical Medicine and Bioethics* 33(1):45–56.
- Beljaars DE, Valckx WJ, Stepan C, Donis J, Lavrijsen JC (2015) Prevalence differences of patients in vegetative state in The Netherlands and Vienna, Austria: a comparison of values and ethics. *The Journal of Head Trauma Rehabilitation* 30(3):E57-60
- Benbadis SR, LaFrance WC Jr, Papandonatos GD, Korabathina K, Lin K, Kraemer HC, NES Treatment Workshop (2009) Interrater reliability of EEG-video monitoring. *Neurology* 73(11):843-846.

Berg M and Mol A (1998) *Differences in medicine: Unraveling practices, techniques, and bodies*. Durham/London: Duke University Press.

Bernacki RE and Block SD (2014) Communication about serious illness care goals: a review and synthesis of best practices. *JAMA internal medicine* 174(12):1994-2003.

Bernard SA, Gray TW, Buist MD, Jones BM, Silvester W, Gutteridge G, Smith K (2002) Treatment of Comatose Survivors of Out-of-Hospital Cardiac Arrest with Induced Hypothermia. *New England Journal of Medicine* 346(8):557–563.
doi:10.1056/NEJMoa003289

Bernheim JL, Vansweevelt T, Annemans L (2013) Medical futility and end-of-life issues in Belgium. In A. Bagheri (ed) *Medical futility: A cross-national study*, 59-83. London: Imperial College Press.

Biggs M (2009) Self-fulfilling Prophecies. In P. Bearman and P. Hedström (eds) *The Oxford handbook of analytical sociology*, 294-314. Oxford: Oxford University Press.

Blandin K and Pepin R (2017) Dementia grief: A theoretical model of a unique grief experience. *Dementia* 16(1):67–78.

Blok V and Lemmens P (2015) The Emerging Concept of Responsible Innovation. Three Reasons Why It Is Questionable and Calls for a Radical Transformation of the Concept of Innovation. In B-J. Koops, I. Oosterlaken, H. Romijn, T.E. Swierstra, J. van den Hoven (eds) *Responsible Innovation 2: Concepts, Approaches, and Applications*, 19–35. Cham: Springer International.

Blom MT, Beesems SG, Homma PC, Zijlstra JA, Hulleman M, van Hoeijen DA, Bardai A, Tijssen JG, Tan HL, Koster RW (2014) Improved survival after out-of-hospital cardiac arrest and use of automated external defibrillators. *Circulation* 130(21):1868-1875.

Boenink M, van Lente H, and Moors E (2016) *Emerging Technologies for Diagnosing Alzheimer's Disease: Innovating with Care*. London: Palgrave Macmillan UK.

Boenink M, van Lente H, and Moors E (2016) Diagnosing Alzheimer's Disease: How to Innovate with Care. In M. Boenink, H. van Lente, and E. Moors (eds) *Emerging Technologies for Diagnosing Alzheimer's Disease: Innovating with Care*, 263-275. London: Palgrave Macmillan Ltd.

Boenink M (2013) The Multiple Practices of Doing 'Ethics in the Laboratory': A Mid-level Perspective. In S. Van der Burg and T.E. Swierstra (eds) *Ethics on the Laboratory Floor*, 57–78. Basingstoke: Palgrave Macmillan.

Boenink M (2012) Debating the desirability of new biomedical technologies: Lessons from the introduction of breast cancer screening in the Netherlands. *Health Care Analysis*, 20(1): 84–102.

Boenink M, Swierstra TE, and Stemmering D (2010) Anticipating the Interaction Between Technology and Morality: A Scenario Study of Experimenting with Humans in Bionanotechnology. *Studies in Ethics, Law, and Technology* 4(2):1941–6008.

Bridges JFP, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA, Johnson FR, Mauskopf J (2011) Conjoint Analysis Applications in Health—a Checklist: A Report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value in Health* 14(4):403–413.

Bro-Jeppesen J, Kjaergaard J, Horsted TI, Wanscher MC, Nielsen SL, Rasmussen LS, Hassager C (2009) The impact of therapeutic hypothermia on neurological function and quality of life after cardiac arrest. *Resuscitation* 80:171–76.

Callahan D (1991) Medical Futility, Medical Necessity: The-Problem-without-a-Name. *The Hastings Center Report* 21(4):30.

Caron-Flinterman JF, Broerse JE, and Bunders JFG (2005) The Experiential Knowledge of Patients: A New Resource for Biomedical Research? *Social Science & Medicine* 60(11):2575–2584.

Cartwright W (1996) Killing and letting die: a defensible distinction. *British Medical Bulletin* 2(2): 354–361.

Chan D, Livingston G, Jones L, Sampson EL (2013) Grief reactions in dementia carers: a systematic review. *International Journal of Geriatric Psychiatry* 28(1):1–17.

Code L (2020) *Epistemic Responsibility*. Albany, NY: State University of New York Press.

Collingridge D (1980) *The Social Control of Technology*. New York: St. Martin's Press.

Conrad P (1992) Medicalization and Social Control. *Annual Review of Sociology* 18: 209–232

Corlett JA (2008) Epistemic Responsibility. *International Journal of Philosophical Studies* 16(2), 179–200.

Cronberg T, Lilja G, Horn J, Kjaergaard J, Wise MP, Pellis T, Hovdenes J, Gasche Y, Åneman A, Ståmmet P, Erlinge D, Friberg H, Hassager C, Kuiper M, Wanscher M, Bosch F, Cranshaw J, Kleger GR, Persson S, Undén J, Walden A, Winkel P, Wetterslev J, Nielsen N; TTM Trial Investigators (2015) Neurologic function and health-related quality of life in patients following targeted temperature management at 33°C versus 36°C after out-of-hospital cardiac arrest: a randomized clinical trial. *JAMA Neurology* 72:634–41.

Dale CM, Sinuff T, Morrison LJ, Golan E, Scales DC (2016) Understanding Early Decisions to Withdraw Life-Sustaining Therapy in Cardiac Arrest Survivors. A Qualitative Investigation. *Annals of the American Thoracic Society* 13(7):1115–1122.

de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA (2015) Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide. *The patient* 8(5):373–384.

de Vreede-Swagemakers JJ, Gorgels AP, Weijenberg MP, Dubois-Arbouw WI, Golombeck B, van Ree JW, Knottnerus A, Wellens HJ (1999) Risk indicators for out-of-hospital cardiac arrest in patients with coronary artery disease. *Journal of clinical epidemiology* 52(7):601–607.

Demertzi A, Ledoux D, Bruno MA, Vanhaudenhuyse A, Gosseries O, Soddu A, Schnakers C, Moonen G, Laureys S. (2011) Attitudes towards end-of-life issues in disorders of consciousness: a European survey. *Journal of neurology* 258(6):1058–65.

Dickel S and Schrape J-F (2017) The Renaissance of Techno-Utopianism as a Challenge for Responsible Innovation. *Journal of Responsible Innovation* 4(2):289–294.

Directorate-General for Research and Innovation (2013) *State of the Innovation Union 2012 – Accelerating Change*. Luxembourg: Publications Office of the European Union.

Dowden B (2020) Fallacies. *The Internet Encyclopedia of Philosophy*, ISSN 2161-0002, <https://iep.utm.edu/fallacy/#Self-FulfillingProphecy>, 2020-11-20.

Dragancea I, Wise MP, Al-Subaie N, Cranshaw J, Friberg H, Glover G, Pellis T, Rylance R, Walden A, Nielsen N, Cronberg T, TTM trial investigators (2017) Protocol-driven neurological prognostication and withdrawal of life-sustaining therapy after cardiac arrest and targeted temperature management. *Resuscitation* 117:50–57.

Dutch Ministry of Health (2020) *Wet op de overeenkomst inzake geneeskundige behandeling (WGBO)* [Medical Treatment Contract Act] [In Dutch]. Burgerlijk wetboek boek 7, art 7:446-468.

- Einav S, Bromiker R, Weiniger CF, Matot I (2011) Mathematical Modeling for Prediction of Survival from Resuscitation Based on Computerized Continuous Capnography: Proof of Concept. *Academic Emergency Medicine* 18(5):468-475.
- Einstein A and Infeld L (1938) *The Evolution of Physics*. Cambridge: Cambridge University Press.
- Endisch C, Westhall E, Kenda M, Streitberger KJ, Kirkegaard H, Stenzel W, Storm C, Ploner CJ, Cronberg T, Friberg H, Englund E, Leithner C (2020) Hypoxic-Ischemic Encephalopathy Evaluated by Brain Autopsy and Neuroprognostication After Cardiac Arrest. *JAMA Neurology* 77(11):1–10.
- Felt U, Fouché R, Miller CA, Smith-Doerr L. (2017) *The handbook of science and technology studies*. Cambridge, MA: The MIT Press.
- Fisher E, Boenink M, van der Burg S, Woodbury N (2012) Responsible Healthcare Innovation: Anticipatory Governance of Nanodiagnostics for Theranostics Medicine. *Expert Review of Molecular Diagnostics* 12(8):857–870. doi:10.1586/erm.12.125
- Fisher E and Guston DH (2008) STIR: Socio-technical integration research. Project Proposal for NSF Award #0849101. Center for Nanotechnology in Society, Arizona State University, Tempe, AZ.
- Fisher E (2007) Ethnographic Invention: Probing the Capacity of Laboratory Decisions. *NanoEthics* 1(2):155–165.
- Fisher E, Mahajan RL, and Mitcham C (2006) Midstream Modulation of Technology: Governance From Within. *Bulletin of Science, Technology & Society* 26(6):485-496.
- Fisher M and Ridley S. (2012) Uncertainty in end-of-life care and shared decision making. *Critical Care Resuscitation* 14(1):81-87.
- Foot P (2002) *Moral Dilemmas: and other topics in moral philosophy*. Oxford: Oxford University Press.
- Førde R (1998) Who is to define the futility of treatment--the patient or the physician? *Tidsskr Nor Lægeforen* 118(17):2652-4.
- Friberg H, Cronberg T, Dünser MW, Duranteau J, Horn J, Oddo M (2015) Survey on current practices for neurological prognostication after cardiac arrest. *Resuscitation* 90:158-162.

Funtowicz SO and Ravetz JR (1993) Science for the Post-Normal Age. *Futures* 25(7):735–755.

Geels FW and Smit WA (2000) Failed technology futures: pitfalls and lessons from a historical survey. *Futures* 32(9-10):867–885.

Geocadin RG, Callaway CW, Fink EL, Golan E, Greer DM, Ko NU, Lang E, Licht DJ, Marino BS, McNair ND, Peberdy MA, Perman SM, Sims DB, Soar J, Sandroni C, American Heart Association Emergency Cardiovascular Care Committee (2019) Standards for Studies of Neurological Prognostication in Comatose Survivors of Cardiac Arrest: A Scientific Statement From the American Heart Association. *Circulation* 140(9):e517-e542.

Geocadin RG, Peberdy MA, and Lazar RM (2012) Poor survival after cardiac arrest resuscitation: a self-fulfilling prophecy or biologic destiny? *Critical care medicine* 40(3):979-980.

Geocadin RG, Buitrago MM, Torbey MT, Chandra-Strobos N, Williams MA, Kaplan PW (2006) Neurologic prognosis and withdrawal of life support after resuscitation from cardiac arrest. *Neurology* 67(1):105.

Gerritsen RT, Kesecioglu J, Kompanje EJO, Meulemans E, Mooi B (2009) *Richtlijn nalaten en staken van behandeling en palliatieve zorg na het staken van behandeling bij volwassen IC patiënten* (Guideline withholding and withdrawal of treatment and palliative care after withdrawal of treatment for adult IC patients) [in Dutch] Utrecht: Nederlandse Vereniging voor Intensive Care

Godin B (2008) *Innovation: The History of a Category*. Working Paper.

Gold B, Puertas L, Davis SP, Metzger A, Yannopoulos D, Oakes DA, Lick CJ, Gillquist DL, Holm SY, Olsen JD, Jain S, Lurie KG (2014) Awakening after cardiac arrest and post resuscitation hypothermia: Are we pulling the plug too early? *Resuscitation* 85(2):211–214.

Gonzalez JM (2019) A Guide to Measuring and Interpreting Attribute Importance. *The patient* 12(3):287-295.

Goodson L and Vassar M (2011) An overview of ethnography in healthcare and medical education research. *Journal of Educational Evaluation for Health Professions* 8:4. doi: 10.3352/jeehp.2011.8.4

Grant AC, Abdel-Baki SG, Weedon J, Arnedo V, Chari G, Koziorynska E, Lushbough C, Maus D, McSween T, Mortati KA, Reznikov A, Omurtag A (2014)

EEG interpretation reliability and interpreter confidence: a large single-center study. *Epilepsy & Behavior* 32:102-7.

Grossestreuer AV, Gaieski DF, Abella BS, Wiebe DJ, Moskowitz A, Ikeda DJ, Haukoos JS, Perman SM (2017) Factors associated with post-arrest withdrawal of life-sustaining therapy. *Resuscitation* 110:114-119.

Grunwald A (2004) The Case of Nanobiotechnology. *EMBO Reports* 5(S1):S32–S36. doi:10.1038/sj.embor.7400223

Grunwald A (2014) Technology Assessment for Responsible Innovation. In J. van den Hoven, N. Doorn, T.E. Swierstra, B.-J. Koops, and H. Romijn (eds) *Responsible Innovation 1: Innovative Solutions for Global Issues*, 15–31. Dordrecht: Springer.

Grunwald A and Achternbosch M (2013) Technology Assessment and Approaches to Early Engagement. In N. Doorn, D. Schuurbiers, I. van de Poel, and M.E. Gorman (eds) *Early Engagement and New Technologies: Opening up the Laboratory*, 15–34. Dordrecht: Springer Netherlands.

Guston DH (2014) Understanding 'Anticipatory Governance'. *Social Studies of Science* 44(2):218–242.

Guston DH, Fisher E, Grunwald A, Owen R, Swierstra TE, and van der Burg S (2014) Responsible Innovation: Motivations for a New Journal. *Journal of Responsible Innovation* 1(1):1–8.

Guston DH (2013) 'Daddy, Can I Have a Puddle Gator?': Creativity, Anticipation, and Responsible Innovation. In R. Owen, J. Bessant, and M. Heintz (eds) *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, 109–118. Chichester: John Wiley & Sons Ltd.

Guston DH and Sarewitz D (2002) Real-time Technology Assessment. *Technology in Culture* 24:93–109.

Hassager C, Nagao K, and Hildick-Smith D (2018) Out-of-hospital cardiac arrest: in-hospital intervention strategies. *The Lancet* 391(10124):989–998.

Haywood K, Whitehead L, Nadkarni VM, Achana F, Beesems S, Böttiger BW, Brooks A, Castrén M, Ong ME, Hazinski MF, Koster RW, Lilja G, Long J, Monsieurs KG, Morley PT, Morrison L, Nichol G, Oriolo V, Saposnik G, Smyth M, Spearpoint K, Williams B, Perkins GD; COSCA Collaborators. (2018) COSCA (Core Outcome Set for Cardiac Arrest) in Adults: An Advisory Statement From the International Liaison Committee on Resuscitation. *Circulation* 29;137(22):e783-e801.

Hofmeijer J, Ruijter BJ, and van Putten MJAM (2019) Reply to: "Early EEG for neurologic prognostication: a self-fulfilling prophecy?" *Annals of Neurology* 86(3):474.

Hofmeijer J and van Putten MJ (2016) EEG in postanoxic coma: Prognostic and diagnostic value. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*, 127(4):2047-2055.

Hofmeijer J, Beernink TMJ, Bosch FH, Beishuizen A, Tjepkema-Cloostermans MC, van Putten MJAM (2015) Early EEG contributes to multimodal outcome prediction of postanoxic coma. *Neurology* 85(2):137–143.

Horn J, Hoedemaekers CWE, Hofmeijer J, Jewbali LSD, Koelman JHTM, de Ruijter W (2019) *Richtlijn Prognose van postanoxisch coma*. Utrecht: Nederlandse Vereniging voor Neurologie. https://richtlijndatabase.nl/richtlijn/prognose_van_postanoxisch_coma

Hsu JW, Madsen CD, and Callaham ML (1996) Quality-of-life and formal functional testing of survivors of out-of-hospital cardiac arrest correlates poorly with traditional neurologic outcome scales. *Annals of emergency medicine* 28(6):597-605.

Husak DN (1979) Killing, letting die and euthanasia. *Journal of Medical Ethics* 5(4):200–202.

Hypothermia after Cardiac Arrest Study Group (2002) Mild Therapeutic Hypothermia to Improve the Neurologic Outcome After Cardiac Arrest. *New England Journal of Medicine* 346(8):549–556.

Jacobs JF, van de Poel I, and Osseweijer P (2010) Sunscreens with Titanium Dioxide (TiO₂) Nano-Particles: A Societal Experiment. *NanoEthics* 4(2):103–113.

Jennett B (2002) *The vegetative state: medical facts, ethical and legal dilemmas*. Cambridge: Cambridge University Press.

Jennett B and Bond M (1975) Assessment of outcome after severe brain damage. A practical Scale. *The Lancet* 305(7905):480-484.

Katz DI, Polyak M, Coughlan D, Nichols M, and Roche A (2009) Natural history of recovery from brain injury after prolonged disorders of consciousness: outcome of patients admitted to inpatient rehabilitation with 1-4 year follow-up. *Progress in Brain Research*, 177:73-88.

Kihlstrom JF (2009) Unconscious Cognition. In W.P. Banks (ed), *Encyclopedia of Consciousness*, 2, 411-421. Oxford: Elsevier.

Kitzinger J and Kitzinger C (2013) The 'window of opportunity' for death after severe brain injury: family experiences. *Sociology of Health and Illness* 35(7):1095-1112.

Ko DN and Blinderman CD (2015) Withholding and withdrawing life-sustaining treatment (including artificial nutrition and hydration) In: N. Cherny, M. Fallon, S. Kaasa, R.K. Portenoy, and C.D. Currow (eds) *Oxford Textbook of Palliative Medicine*. Oxford: Oxford University Press.

Kudenchuk PJ, Sandroni C, Drinhaus HR, Böttiger BW, Cariou A, Sunde K, Dworschak M, Taccone FS, Deye N, Friberg H, Laureys S, Ledoux D, Oddo M, Legriel S, Hantson P, Diehl J-L, Laterre P-F, (2015) Breakthrough in cardiac arrest: reports from the 4th Paris International Conference. *Annals of Intensive Care* 5(1):22-45.

Kuhlmann BS and Rip A (2014) The Challenge of Addressing Grand Challenges: a think piece on how innovation can be driven towards the "Grand Challenges" as defined under the prospective European Union Framework Programme Horizon 2020. Report to ERIAB. Brussels: European Commission.

La Shure C (2005) "What is liminality?" <http://www.liminality.org/about/whatisliminality/>.

Lane H (2005) Ethnicity, Ethics, and the Deaf-World. *Journal of Deaf Studies and Deaf Education* 10(3):291–310.

Lane H, Pillard R and French M (2000) Origins of the American Deaf-World: Assimilating and differentiating societies and their relation to genetic patterning. *Sign Language Studies* 1, new series, 17–44.

Lantos JD, Singer PA, Walker RM, Gramelspacher GP, Shapiro GR, Sanchez-Gonzalez MA, Stocking CB, Miles SH, Siegler M (1989) The illusion of futility in clinical practice. *The American Journal of Medicine* 87(1):81-4.

Latour B (1996) *Aramis, or the Love of Technology*. Cambridge, MA: Harvard University Press.

Latour B (1987) *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, MA: Harvard University Press.

Law J and Singleton V (2000) Performing Technology's Stories: On Social Constructivism, Performance, and Performativity. *Technology and Culture* 41(4):765-775.

Levin PD and Sprung CL (2005) Withdrawing and withholding life-sustaining therapies are not the same. *Critical care* 9(3):230–232.

Licoppe C (2010) The “Performative Turn” in Science and Technology Studies. *Journal of Cultural Economy* 3(2): 181–188.

Lucivero F, Swierstra TE, and Boenink M (2011) Assessing Expectations: Towards a Toolbox for an Ethics of Emerging Technologies. *NanoEthics* 5(2):129–141.

Maciel CB, Barden MM, and Greer DM (2017) Neurologic Recovery After Cardiac Arrest: a Multifaceted Puzzle Requiring Comprehensive Coordinated Care. *Current treatment options in cardiovascular medicine* 19(7):52.

MacKenzie DA, Muniesa F, and Siu L (2007) *Do economists make markets? On the performativity of economics*. Princeton, NJ: Princeton University Press.

Macnaghten P, Owen R, Stilgoe J, Wynne B, Azevedo A, de Campos A, Chilvers J, Dagnino R, di Giulio G, Frow E, Garvey B, Groves C, Hartley S, Knobel M, Kobayashi E, Lehtonen M, Lezaun J, Mello L, Monteiro M, Pamplona da Costa J, Rigolin C, Rondani B, Staykova M, Taddei R, Till C, Tyfield D, Wilford S, Velho L (2014) Responsible Innovation Across Borders: Tensions, Paradoxes and Possibilities. *Journal of Responsible Innovation* 1(2):191–199.

Mertens M (2018) Liminal innovation practices: questioning three common assumptions in responsible innovation. *Journal of Responsible Innovation* 5(3):280-298.

Mertens M, King OC, van Putten MJAM, Boenink M (2021) Can we learn from hidden mistakes? Self-fulfilling prophecy and Responsible Neuroprognostic Innovation. *Journal of Medical Ethics*.

Mertens M, van Til JA, Bouwers-Beens E, Boenink M (2021) Chasing certainty After Cardiac Arrest: Can a technological innovation solve a moral dilemma? *Neuroethics*

Merton RK (1948) The self-fulfilling prophecy. *The Antioch Review* 8(2):193-210.
Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L (2016) The ethics of algorithms: Mapping the debate. *Big Data & Society* 3(2):205395171667967.

Mohindra RK (2007) Medical futility: a conceptual model. *Journal of Medical Ethics* 33:71–75.

Moratti S (2010) The Englaro Case: withdrawal of treatment from a patient in a permanent vegetative state in Italy. *Cambridge quarterly of healthcare ethics* 19(3):372-80.

Moulaert VRMP (2014) *Life after survival of cardiac arrest. The brain is the heart of the matter*. Enschede: Gildeprint.

Moulaert VR, Verbunt JA, van Heugten CM, Wade DT (2009) Cognitive impairments in survivors of out-of-hospital cardiac arrest: a systematic review. *Resuscitation* 80(3):297-305.

Mulder M, Gibbs HG, Smith SW, Dhaliwal R, Scott NL, Sprenkle MD, Geocadin RG (2014) Awakening and withdrawal of life-sustaining treatment in cardiac arrest survivors treated with therapeutic hypothermia*. *Critical Care Medicine* 42(12):2493-2499.

Myat A, Song K-J, and Rea T (2018) Out-of-hospital cardiac arrest: current concepts. *The Lancet* 391(10124):970-979.

Naeije G, Depondt C, Meeus C, Korpak K, Pepersack T, Legros B (2014) EEG Patterns Compatible with Nonconvulsive Status Epilepticus are Common in Elderly Patients with Delirium: A Prospective Study with Continuous EEG Monitoring. *Epilepsy & Behavior* 36:18–21.

Nagaraj SB, Tjepkema-Cloostermans MC, Ruijter BJ, Hofmeijer J, van Putten MJAM (2018) The revised Cerebral Recovery Index improves predictions of neurological outcome after cardiac arrest. *Clinical neurophysiology* 129(12):2557-2566.

National Academies of Sciences, Engineering, and Medicine (2017) *The Promise of Assistive Technology to Enhance Activity and Work Participation*. Washington, DC: The National Academies Press.

Nielsen N, Wetterslev J, Cronberg T, Erlinge D, Gasche Y, Hassager C, Horn J, Hovdenes J, Kjaergaard J, Kuiper M, Pellis T, Stammet P, Wanscher M, Wise MP, Åneman A, Al-Subaie N, Boesgaard S, Bro-Jeppesen J, Brunetti I, Bugge JF, Hingston CD, Juffermans NP, Koopmans M, Køber L, Langørgen J, Lilja G, Møller JE, Rundgren M, Rylander C, Smid O, Werer C, Winkel P, Friberg H; TTM Trial Investigators (2013) Targeted Temperature Management at 33°C Versus 36°C After Cardiac Arrest. *New England Journal of Medicine* 369(23):2197–2206.

Nolan JP, Soar J, Cariou A, Cronberg T, Moulaert VR, Deakin CD, Bottiger BW, Friberg H, Sunde K, Sandroni C (2015) European Resuscitation Council and European Society of Intensive Care Medicine 2015 guidelines for post-resuscitation care. *Intensive care medicine* 41(12):2039-2056.

Nolan JP, Sandroni C, Böttiger BW, Cariou A, Cronberg T, Friberg H, Genbrugge C, Haywood K, Lilja G, Moulaert VRM, Nikolaou N, Olasveengen TM, Skrifvars

MB, Taccone F, Soar J. (2021) European Resuscitation Council and European Society of Intensive Care Medicine guidelines 2021: post-resuscitation care. *Intensive Care Medicine* 47(4):369-421.

Nordmann A (2014) Responsible Innovation, the Art and Craft of Anticipation. *Journal of Responsible Innovation* 1(1):87–98.

Núñez Partido JP (2012) Unconscious Learning. In: N.M. Seel (ed) *Encyclopedia of the Sciences of Learning*. Springer, Boston, MA.

Oh SH, Park KN, Shon YM, Kim YM, Kim HJ, Youn CS, Kim SH, Choi SP, Kim SC (2015) Continuous Amplitude-Integrated Electroencephalographic Monitoring Is a Useful Prognostic Tool for Hypothermia-Treated Cardiac Arrest Patients. *Circulation* 132(12):1094–1103.

Owen R, Macnaghten P, and Stilgoe J (2012) Responsible Research and Innovation: From Science in Society to Science for Society, with Society. *Science and Public Policy* 39(6):751–760.

Parthasarathy S (2012) *Building genetic medicine: Breast cancer, technology, and the comparative politics of health care*. Cambridge MA: MIT Press.

Parthasarathy S (2005) Architectures of genetic medicine: comparing genetic testing for breast cancer in the USA and the UK. *Social Studies of Science* 35(1): 5–40.

Pellegrino ED (2005) Futility in Medical Decisions: The Word and the Concept. *HEC Forum* 17(4):308–318.

Pickering A (1995) Beyond constraint: The temporality of practice and the historicity of knowledge. In: J.Z. Buchwald (ed) *Scientific practice: Theories and stories of doing physics*, 42–55. Chicago: University of Chicago Press.

Pols J (2014) Knowing Patients: Turning Patient Knowledge Into Science. *Science, Technology, & Human Values* 39(1):73–97.

Quill TE, Holloway R (2011) Time-limited trials near the end of life. *Journal of the American Medical Association* 306:1483–1484.

Rachels J (2001) Killing and Letting Die. In: L.R. Becker, M. Becker, and C. Becker (eds) *Encyclopedia of Ethics*, 2nd edition, 947-950. New York: Routledge.

Raina KD, Callaway C, Rittenberger JC, Holm MB (2008) Neurological and functional status following cardiac arrest: method and tool utility. *Resuscitation* 79(2):249-256.

Reed Johnson F, Lancsar E, Marshall D, Kilambi V, Mühlbacher A, Regier DA, Bresnahan BW, Kanninen B, Bridges JFP (2013) Constructing Experimental Designs for Discrete-Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value in Health* 16(1):3-13.

Robertson-Steel I (2006) Evolution of triage systems. *Emergency Medicine Journal* 23(2):154-5.

Rotolo D, Hicks D, and Martin BR (2015) What is an Emerging Technology? *Research Policy* 44(10):1827–1843.

Roulstone A (2016) *Disability and Technology: An Interdisciplinary and International Approach*. London: Palgrave Macmillan.

Rosenthal R and Rubin DB (1978) Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences* 1(3):377-386.

Rossetti AO, Rabinstein AA, Oddo M (2016) Neurological prognostication of outcome in patients in com after cardiac arrest. *The Lancet Neurology* 15(6):597-609.

(Royal Dutch Medical Association) KNMG / Legemaate J (2005) *Begrippen en zorgvuldigheidseisen rond het levenseinde (Terms and due diligence around end of life)* [in Dutch]. Utrecht: (Royal Dutch Medical Association) KNMG.

(Royal Dutch Medical Association) KNMG / de Jong EC, Legemaate J, Dillmann RJM (1997) Commissie Aanvaardbaarheid Levensbeëindigend handelen. *Medisch handelen rond het levenseinde bij wilsonbekwame patiënten (Medical end-of-life practice for incompetent patients)* [in Dutch]. Houten/Diegem, the Netherlands: Bohn Stafleu Van Loghum.

Ruijter BJ, Tjepkema-Cloostermans MC, Tromp SC, Bergh WM, Foudraine NA, Kornips FHM, Drost G, Scholten E, Bosch FH, Beishuizen A, van Putten MJAM, Hofmeijer J (2019) Early EEG for outcome prediction of postanoxic coma: a prospective cohort study. *Annals of Neurology* 86(2):203–214.
doi:10.1002/ana.25518

Ruijter BJ, Hofmeijer J, Tjepkema-Cloostermans MC, van Putten M (2018) The prognostic value of discontinuous EEG patterns in postanoxic coma. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*, 129(8):1534-1543.

Ryan M (2004) Discrete choice experiments in health care. *The BMJ* 328(7436):360-361.

Rydvall A and Lynöe N (2008) Withholding and withdrawing life-sustaining treatment: a comparative study of the ethical reasoning of physicians and the general public. *Critical care* 12(1):R13.

Sandroni C, D'Arrigo S, and Nolan JP (2018) Prognostication after cardiac arrest. *Critical Care* 22(1):150.

Sandroni C and Nolan JP (2015) Neuroprognostication after cardiac arrest in Europe: New timings and standards. *Resuscitation* 90: A4–A5.

Sandroni C, Cariou A, Cavallaro F, Cronberg T, Friberg H, Hoedemaekers C, Horn J, Nolan JP, Rossetti AO, Soar J (2014) Prognostication in comatose survivors of cardiac arrest: An advisory statement from the European Resuscitation Council and the European Society of Intensive Care Medicine. *Intensive Care Medicine* 40(12):1816–1831.

Schembs L, Ruhfass M, Racine E, Jox RJ, Bender A, Rosenfelder M, Kuehlmeier K (2020) How Does Functional Neurodiagnostics Inform Surrogate Decision-Making for Patients with Disorders of Consciousness? A Qualitative Interview Study with Patients' Next of Kin. *Neuroethics*.

Schenker Y, Tiver GA, Hong SY, White DB (2013) Discussion of treatment trials in intensive care. *Journal of Critical Care* 28:862–869.

Schneiderman LJ (1990) Medical Futility: Its Meaning and Ethical Implications. *Annals of Internal Medicine*, 112(12):949.

Schneiderman L.J (2011) Defining medical futility and improving medical care. *Bioethical Inquiry* 8:123–131.

Schroeder D and Ladikas M (2015) Towards Principled Responsible Research and Innovation: Employing the Difference Principle in Funding Decisions. *Journal of Responsible Innovation* 2(2):169–183.

Scollo-Lavizzari G and Bassetti C (1987) Prognostic Value of EEG in Post-Anoxic Coma After Cardiac Arrest. *European Neurology* 26(3):161–170.

Selin C (2008) The Sociology of the Future: Tracing Stories of Technology and Time. *Sociology Compass* 2(6):1878–1895.

Silva S, Kenney M. (2019). Algorithms, platforms, and ethnic bias. *Communications of the ACM* 62(11):37–39.

Simons R-JP (2012) Unconscious and Conscious Learning. *Encyclopedia of the Sciences of Learning* 3363–3366.

Shiffrin SV (2012) Harm And Its Moral Significance. *Legal Theory* 18(03):357–398.

Smith AK, White DB, and Arnold RM (2013) Uncertainty—The Other Side of Prognosis. *New England Journal of Medicine* 368(26):2448–2450.

Sokol DK (2009) The slipperiness of futility. *BMJ*, 338, b2222

Sondag L, Ruijter BJ, Tjepkema-Cloostermans MC, Beishuizen A, Bosch FH, van Til JA, van Putten M, Hofmeijer J (2017) Early EEG for outcome prediction of postanoxic coma: prospective cohort study with cost-minimization analysis. *Critical care* 21(1):111.

Spalletti M, Carrai R, Scarpino M, Cossu C, Ammannati A, Ciapetti M, Tadini Buoninsegni L, Peris A, Valente S, Grippo A, Amantini A (2016) Single Electroencephalographic Patterns as Specific and Time-dependent Indicators of Good and Poor Outcome After Cardiac Arrest. *Clinical Neurophysiology* 127(7):2610–2617.

Steinberg A and Sprung CL (2007) The dying patient act, 2005: Israeli innovative legislation. *Israel Medical Association Journal* 9(7):550-552

Steinberg A, Callaway CW, Arnold RM, Cronberg T, Naito H, Dadon K, Chae MK, Elmer J (2019) Prognostication after cardiac arrest: Results of an international, multi-professional survey. *Resuscitation* 138:190-197.

Stilgoe J, Owen R, and Macnaghten P (2013) Developing a Framework for Responsible Innovation. *Research Policy* 42(9):1568–1580.

Swierstra TE and Keulartz J (2011) Obesity in 2020: Three Scenarios on Techno-Socio-Ethical Co-Evolution. I. In M. Korthals (ed) *Genomics, Obesity and the Struggle Over Responsibilities*, 97–112. Dordrecht: Springer Academic.

Swierstra TE, Stermerding D, and Boenink M (2009) Exploring Techno-Moral Change: The Case of the Obesity Pill. In P. Sollie and M. Düwell (eds) *Evaluating New Technologies: Methodological Problems for the Ethical Assessment of Technology Developments*, 119–138. Dordrecht: Springer Netherlands.

Taccone F, Cronberg T, Friberg H, Greer D, Horn J, Oddo M, Scolletta S, Vincent JL (2014) How to assess prognosis after cardiac arrest and therapeutic hypothermia. *Critical Care* 18(1):202.

Thomson J (1976) Killing, Letting Die, and the Trolley Problem. *The Monist* 59(2):204–217.

Tjepkema-Cloostermans MC, da Silva Lourenço C, Ruijter BJ, Tromp SC, Drost G, Kornips FHM, Beishuizen A, Bosch FH, Hofmeijer J, van Putten MJAM (2019) Outcome Prediction in Postanoxic Coma With Deep Learning. *Critical Care Medicine* 47(10):1424–1432

Tjepkema-Cloostermans MC, de Carvalho RCV, van Putten MJAM (2018) Deep learning for detection of focal epileptiform discharges from scalp EEG recordings. *Clinical Neurophysiology* 129(10):2191–2196.

Tjepkema-Cloostermans MC, Hofmeijer J, Beishuizen A, Hom HW, Blans MJ, Bosch FH, van Putten MJAM (2017) Cerebral recovery index: Reliable help for prediction of neurologic outcome after cardiac arrest. *Critical Care Medicine* 45(8):e789–e797.

Tjepkema-Cloostermans MC, Hofmeijer J, Trof R.J, Blans M.J, Beishuizen A and van Putten MJAM (2015) Electroencephalogram predicts outcome in patients with postanoxic coma during mild therapeutic hypothermia. *Critical Care Medicine* 43(1): 159–167.

Tjepkema-Cloostermans MC, van Meulen FB, Meinsma G and van Putten MJAM (2013) A Cerebral Recovery Index (CRI) for early prognosis in patients after cardiac arrest. *Critical Care*, 17(5): R252.

Tong JT, Eyngorn I, Mlynash M, Albers GW, Hirsch KG (2016) Functional Neurologic Outcomes Change Over the First 6 Months After Cardiac Arrest. *Critical Care Medicine* 44(12):e1202–e1207.

Tooley M (1980) An Irrelevant Consideration: Killing versus Letting Die. In: Steinbock B (ed) *Killing and Letting Die*, 56–62. Englewood Cliffs: Prentice-Hall.

Turner VW (1967) *Between and Between: The Liminal Period in Rites de Passage*. In *The Forest of Symbols: Aspects of Ndembu Ritual*. Ithaca: Cornell University Press.

Turner VW (1969) *Liminality and Communitas*. In *The Ritual Process: Structure and Anti-Structure*, 94–130. Chicago: Aldine.

United States, National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1978) *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research by the National Commission for the*

Protection of Human Subjects of Biomedical and Behavioral Research, DHEW Publication (OS) 78-0012. Washington, D.C.: Government Printing Office.

van de Poel I (2009) The Introduction of Nanotechnology as a Societal Experiment. In S.S. Arnaldi, A. Lorenzet, and F. Russo (eds) *Technoscience in Progress. Managing the Uncertainty of Nanotechnology*, edited by, 129–142. Amsterdam: IOS Press.

van de Poel I (2011) Nuclear Energy as a Social Experiment. *Ethics, Policy & Environment* 14 (3):285–290.

van de Poel I (2013) Why New Technologies Should be Conceived as Social Experiments. *Ethics, Policy & Environment* 16(3):352–355.

van der Burg S and Swierstra TE (2013) *Ethics on the laboratory floor*. London: Palgrave Macmillan.

van der Kooi AW, Leijten FS, van der Wekken RJ, Slooter AJ (2012) What Are the Opportunities for EEG-based Monitoring of Delirium in the ICU? *The Journal of Neuropsychiatry and Clinical Neurosciences* 24(4):472–477

van Erp W (2020) *The unresponsive wakefulness syndrome in the Netherlands: outcomes from a vicious circle*. Utrecht: Libertas Pascal.

Van Putten MJAM, Jansen C, Tjepkema-Cloostermans MC, Beernink TMJ, Koot R, Bosch F, Beishuizen A, Hofmeijer J (2019) Postmortem histopathology of electroencephalography and evoked potentials in postanoxic coma. *Resuscitation* 134:26–32.

Van Putten MJAM, Hofmeijer J, Ruijter BJ, Tjepkema-Cloostermans MC (2017) Deep Learning for outcome prediction of postanoxic coma. In: H. Eskola, O. Väisänen, J. Viik, J/ Hyttinen (eds) *EMBEC & NBC 2017*. IFMBE Proceedings, vol 65. Singapore: Springer.

Van Til J, Bouwers-Beens E, Mertens M, Boenink M, Groothuis K, Hofmeijer J (2021) Prognostication of patients in coma after cardiac arrest: public perspectives. Manuscript submitted for publication.

Veatch RM (2013) Medical futility and end-of-life issues in Belgium. In A. Bagheri (ed) *Medical futility: A cross-national study*, 9–33. London: Imperial College Press.

Verbeek PPCC (2008) Morality in design: Design ethics and the morality of technological artifacts. In Vermaas PE, Kroes P, Light A, Moore S (eds) *Philosophy and Design*, 91–103. Dordrecht: Springer.

Verbeek PPCC (2006) Materializing Morality—design ethics and technological mediation. *Science, Technology and Human Values* 31(3):361–380.

Verbeek PPCC (2005) *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. University Park, PA: The Pennsylvania State University Press.

Vincent JL (2005) Withdrawing may be preferable to withholding. *Critical Care* 9:226–229.

Vink EE, Azoulay E, Caplan A, Kompanje EJO, Bakker J (2018) Time-limited trial of intensive care treatment: an overview of current literature. *Intensive Care Medicine* 44(9):1369–177.

von Schomberg R (2013) A Vision of Responsible Research and Innovation. In R. Owen, J. Bessant, and M. Heintz (eds) *Responsible Innovation. Managing the Responsible Emergence of Science and Innovation in Society*, 51–74. London: John Wiley.

von Schomberg R (2011) Prospects for Technology Assessment in a Framework of Responsible Research and Innovation. In M. Beecroft and R. Dusseldorp (eds) *Technikfolgen Abschätzen Lehren: Bildungspotenziale Transdisziplinärer Methoden*, 39–61. Wiesbaden: Springer VS.

Wallenburg I, Quartz J, Bal R (2019) Making Hospitals Governable: Performativity and Institutional Work in Ranking Practices. *Administration & Society* 51(4):637–663.

Welie JV and Ten Have HA (2014) The ethics of forgoing life-sustaining treatment: theoretical considerations and clinical decision making. *Multidisciplinary respiratory medicine*, 9(1)14.

Wender BA, Foley RW, Hottle TA, Sadowski J, Prado-Lopez V, Eisenberg DA, Laurin L, Seager TP (2014) Anticipatory Life-cycle Assessment for Responsible Research and Innovation. *Journal of Responsible Innovation* 1(2):200–207.

Westhall E, Rossetti AO, van Rootselaar AF, Wesenberg Kjaer T, Horn J, Ullén S, Friberg H, Nielsen N, Rosén I, Åneman A, Erlinge D, Gasche Y, Hassager C, Hovdenes J, Kjaergaard J, Kuiper M, Pellis T, Stammet P, Wanscher M, Wetterslev J, Wise MP, Cronberg T, TTM-trial investigators (2016) Standardized EEG interpretation accurately predicts prognosis after cardiac arrest. *Neurology* 86:1482–1490.

White K (2017) *An introduction to the sociology of health and illness*. London: Sage Publications.

Wilkinson D (2009) The self-fulfilling prophecy in intensive care. *Theoretical medicine and bioethics* 30(6):401-410.

Wilkinson D (2011) The window of opportunity for treatment withdrawal. *Archives of Pediatrics & Adolescent Medicine* 165(3):211-5.

Wilkinson D (2013) *Death or Disability? The 'Carmentis Machine' and decision-making for critically ill children*. Oxford: Oxford University Press.

Worster A and Carpenter C (2008) Incorporation bias in studies of diagnostic tests: how to avoid being biased about bias. *Canadian Journal of Emergency Medicine* 10(2):174-175.

Yan S, Gan Y, Jiang N, Wang R, Chen Y, Luo Z, Zong Q, Chen S, Chuanzhu Lv (2020) The global survival rate among adult out-of-hospital cardiac arrest patients who received cardiopulmonary resuscitation: a systematic review and meta-analysis. *Critical Care* 24(1):61.

Youngner SJ (1988) Who Defines Futility? *The Journal of the American Medical Association* 260(14):2094.

Zandbergen EGJ, van Dijk JG, Hoedemakers C, Horn J, Hijdra A, Kuiper MA, van der Naalt J, Zandstra DF (2011) *Guideline Prognosis of Postanoxic Coma*. [In Dutch] *Richtlijn Prognose van Post-Anoxisch Coma*. Nederlandse Vereniging voor Neurologie, Nederlandse Vereniging voor Intensive Care.

Zandbergen EGJ, Koelman JHTM, de Haan RJ, Hijdra A (2006) SSEPs and prognosis in postanoxic coma: only short or also long latency responses? *Neurology* 67(4):583-586.

Zandbergen EG, Hijdra A, Koelman JH, Hart AA, Vos PE, Verbeek MM, de Haan RJ (2006) Prediction of poor outcome within the first 3 days of postanoxic coma. *Neurology* 66(1):62-68.

Zandbergen EG, de Haan RJ, Koelman JH, Hijdra A (2000) Prediction of poor outcome in anoxic-ischemic coma. *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society* 17(5):498-501.

Zandbergen EGJ, de Haan RJ, Stoutenbeek CP, Koelman JHTM, Hijdra A (1998) Systematic review of early prediction of poor outcome in anoxic-ischaemic coma. *The Lancet* 352(9143):1808-1812.

Zola IK (1972) Medicine as an Institution of Social Control. *The Sociological Review* 20(4):487-504.

Acknowledgements

"Nothing of me is original. I am the combined effort of everyone I've ever known."

I start my words of gratitude with this quote by Chuck Palahniuk to remind everyone that if you've touched my life, I carry you with me. Even if your impact on my life may seem irrelevant to this work, we can never be sure how it reverberated through time and space. So, whether or not we've ever met, thank you for your contribution, whatever it was.

Mom. Words will never do justice to the gratitude I would like to extend to you but one of the countless things you have taught me is that unexpressed gratitude is just the same as no gratitude at all. Of course, I owe you my life, quite literally, but even the *kind* of life I live—and it's been an extraordinary one so far—is to such a large extent thanks to you. Thank you for your endless support, your relentless belief in me, and your increasingly growing affection. In these last years, you managed to simultaneously let go of me and be more present than ever. Most important of all, you've been an inspiring example, a role model. Your own transformation has been one of the greatest inspirations of all and I've been practicing religiously, every single day. Thank you for building a foundation *for* me and, later, *with* me.

Wouter. You're not just part of my foundation, the place that I can always run to, the person that feels like home. You're also one of my favorite people in the world and I doubt this will ever change. I don't know anyone with your heart. Your tendency to give and your natural attention for others is admirable. I am so grateful to have you as my big brother. I love you.

I loved a younger brother too. I remember mom saying that, considering the brain damage you had suffered, she was happy when attempts to resuscitate you failed. While you didn't get much chance to experience the sister I would have been to you, I wish you could know that I carry you with me everywhere I go. An eternal tear in my soul, you keep me compassionate and forgiving. You are the constant reminder of what really matters in life and you taught me early on that there is beauty in grief. It has made me value even the worst parts of life.

Owen. I think everyone knows this work would not have been the same without you. Your co-authorship, valuable as it is, is but a fraction of your contribution. Working with you strengthened this work's combination of empirical and conceptual analysis and the result of our efforts was always more than the sum of its parts. I feel incalculably lucky for all the hours, days, and weeks we spent discussing each other's work and, ultimately, *our* work. I am grateful for all the feedback I was able to give to and receive from you, both in person and in writing. Whether sitting

side by side or for hours on the phone, regardless of the circumstances (and damn, were those challenging at times!), we would work together to get the sentences just right. Our writing brought us together long before we collaborated but collaborating with you taught me a lot. I discovered what I really like about analytic philosophy and, in the end, also what I dislike about it. I learned to appreciate my own ability for critical and analytical thought and you helped me identify and acknowledge my passion for epistemology; a love I believe we share. Finally, I can see the good in that I had to bring this work to an end without you. A PhD, after all, is proof of your ability to work independently. Nevertheless, I owe a lot of my growth to you, both professionally as well as personally. For all that, I thank you.

Melis. Our friendship was foundational throughout the PhD-process and I know I found in you a friend for life. I am almost grateful for the challenges we both faced if only because it taught us how much we can rely on one another and ultimately, how much we can rely on ourselves. A true friend helps you to help yourself. We established such a true friendship and I am thrilled to see where all these years of independent struggle have led both of us. I am so proud of you and I am honored to be your friend. Thanks for the unfaltering support you've given me while wrapping up. You are the best paronymph someone could wish for. Thank you, also, for introducing me to Duygu and Dennis.

Alongside Melis' colorful character... Olya, Agata, Bas, Jonne, Sophie, David, and Cathrin too, thank you all for making the struggle so much more bearable. I will never think back on my PhD-life without thinking of you guys. I love that we were all in it together. For those of us who still have to finish, I'm looking forward to your graduation and I'll be cheering you on until then.

Vanja. Thank you for the big sister that you have been during these past 6 years. I am so grateful for how our relationship evolved and I can't wait to be done with this and see you and the kids more often than I have. Love you.

Sarah... my kid sister... I thought I wouldn't know where you'd be on this day but where you are now, you cannot leave. Know that you are here in my heart. You and that insane-making life of yours definitely had an impact on this PhD trajectory of mine. I hope to hand you this book someday, if not today, and maybe you'll understand why I had to take a breather and was not able to do more than I did.

Monique. You've been a pillar throughout this journey and you've helped me grow way beyond the professional difficulties and personal challenges I encountered during this academic journey. And were there many of those... I owe you so much, including my graduation. Perhaps my sanity too. Our conversations are hidden behind the pages of this book because, without your support, I doubt there would be any book at all.

Thank you, Daphne, for all the guidance you offered during several significant events in the first years of my PhD-trajectory and for introducing me to Monique once your counseling was completed.

Johan de Jong. Thank you for all your support and encouragement. I hope you will continue to guide many PhDs into healthier trajectories.

Marianne, my daily supervisor and much deserved first promoter. I've rarely had a working relationship with someone so different from myself. I'm happy that, over the years, we found ways to make our differences work for us. In many ways, I was lucky. At the start, you gave me the kind of freedom that I needed in order to be able to continue. You were always careful, always patient, never carried away quickly... like I said, very different from me. You had a sobering effect on me and sometimes that halted me in my enthusiasm to the point that I felt disheartened. While there may have been times in the beginning when you pushed too little, carefully considering my circumstances, there were definitely times towards the end, once you learned how resilient I could sometimes be, when you pushed too hard. I say this with a smile. Especially since, sometimes, this trial and error of finding a balance and our combined stubbornness led to remarkable results. Ultimately, I think we can both be proud of what we accomplished. You have definitely put in the work. Thank you.

Sven. I cannot thank you enough for all you have contributed to my journey as a starting academic. First of all, I don't know anyone, besides Marianne, who read my work as much and as carefully as you have. But aside from your comments, feedback, editing, proofreading, you've taught so many of us how to survive the challenges and rejections of academic journal publishing. You've also been an essential support to several of my close friends at times when they were in dire need of it, nearly succumbing to the pressure and hostility sometimes too common in PhD-life. Your generous attitude towards those in positions below you gave us all hope that academia does not have to be quite as toxic as we sometimes experienced it to be.

Maren Behrens, thank you for your input and guidance both during my master's and PhD. I will always appreciate your candid, unabridged feedback.

Philip Brey. In 5 years, I've learned two things from you. How to maximise results with minimal effort (including the limitations of such strategies) and the art of delegation (including its pitfalls). I say "learned" but I doubt I'll ever excel at either. While often confusing and sometimes conflicting, the combination of you and Marianne, at times, gave a surprisingly valuable result, benefitting the quality of the work overall. Even if your input was extremely limited, I'd be lying if I

didn't recognize you've made a couple of contributions that were foundationally significant. As said, minimal effort, maximum result.

Ciano. I feel like both of us simultaneously started a new journey. You as head of the department and I as a doctoral student in that department. While you were often limited by external forces, you always made me feel like you had my back. Whether you were desperately trying to get me an office and cringed every time you saw me move my stuff around on a chair during my first months at the department, whether you were making sure I was getting professional support and a substantial extension, or just by lending a listening ear and offering words of support and encouragement. You made sure I knew your door was always open. Thank you.

Margoth, Leon, and Sage, thanks for picking up the torch after the previous generation of PhD's had left it cold. Karolina, Philip, Sean, Isaac, Rosalie, Haleh, Jesse, Shaked, Jana, Masinde, ... I wish you all good luck and hope you either finish your dissertation or happily walk away towards other exciting endeavors. The same counts for all the PhD's at 4TU.Ethics and OZSW. Patyt too, of course, but you deserve your own special dedication. For you, special! How lucky were we that we got stuck in quarantine together? Thank you for all the laughter (and the crying), the stories, the dreams, the songs, the movies, the eating and the drinking, and the yoga ... It was fun!

Mark, you are the funniest office buddy someone could wish for but you are also a rock. 'Just do it! It's not the *fooking* Mona Lisa.' Thank you for your ongoing encouragement and support.

Kritika, if anyone has offered me support in those final years, it is you. Thank you for your collegial spirit, your relentless encouragement, and unwaveringly loyal friendship and love. It was so good and refreshing to have someone in the trenches with me who shared my recognition of other people's "real" problems, enough to continue to feel fortunate and grateful, despite the troubles we faced. Thank you also for all the hours of writing I got done in your company both at writing retreats and online. The daily writing sessions with you and Conrad were a blessing during the office-less months. Thanks, Conrad, also for reading my new conclusion at the very last minute.

Patrick Taylor Smith. In life, we need pillars and bridges but also extenders. You pushed me way out of my comfort zone and forced me to assert myself in opposition at a time when I desperately needed to rediscover the boundaries that defined me.

Robert, it's been fascinating to see you balance between offering support and pushing me to stand tall, yet never allowing me to lean on you in ways that would make me dependent. Thank you for letting me walk with you.

Lori Bruce. You were and will remain an inspiration! I have learned so much from you and there's still so much left to learn. Thank you for being a mentor and showing me how to stand tall in the face of adversity and exemplify how one can always do more. Thank you also for giving me the opportunity to teach at your wonderful bioethics program and allow all of us teaching staff to develop important curricula. As a pre-Lori alumna, I've seen that program transform in your hands and it's left me excited and proud to be a part of.

I also want to thank Stephen Latham, Carol Pollard, and the whole Yale faculty at the Interdisciplinary Center for Bioethics. Shizuko, Zohar, Dan, Evie (K), AJ, Sheena, Brian, Ramona, Jennifer, Laure, Santiago, Santa, Karmele, Shelly, Stephen, Matthew, Natalie, Susan, Evie (M), Agatha, Joe, Jeff, Ruth, and especially Ron... Some of you I'm lucky enough to count as friends but all of you are exquisite colleagues and I feel fortunate to work with you. Thank you for your contributions, great and small. Many thanks also to my students. I love every single class with you.

Thanks to the people at Hersenstichting, Twente Medical Systems International, and Clinical Science Systems for helping fund this project. Of course, without the support of the Dutch Foundation for Scientific Research (NWO) there would likely be no project to speak of. Thank you.

My greatest debt goes to the patients and their families who so generously opened up their lives, and sometimes, the process of their death for this research. Ingrid, Lars, Jorn, Jolande, Gerard, Jannie, Niek, Riet, Aris, Anja, Jenna, ... I thank you for your courage and contributions. You have been wonderful. Harry and Henk, I very much regret not having met you before your injury.

I also thank all of my other respondents and the experts who gave their input at the valorisation panels and workshops.

Special thanks to the medical staff at Medisch Spectrum Twente, Rijstate, UMC Utrecht, Yale-New Haven Hospital, Tel Aviv Sourasky Medical Center, Soroka Beersheba, Shaare Zedek Jerusalem, and Herzog Medical Center Jerusalem.

Thanks to Carolina Maciel, Teddy Youn, David Greer, and Aneeta Saxena for patiently guiding me into the world of postanoxic coma.

Very special thanks to Michel van Putten, Jeannette Hofmeijer, and Lawrence Hirsch for opening up their practices, introducing me to their colleagues and sharing their concerns.

Michel. You've been a true ally from beginning to finish. Thank you so much for your years of guidance and friendship.

Jeannette. It's always good to have the harshest critic on your team. Although we may come at the same problem from very different angles, I very much appreciate how your perspective and input enriched the work.

Janine, Marieke, and Eline. Thank you for the valuable exchange, your input, collaboration, and co-authorship.

Willemijn. Thank you for the cheerful enthusiasm you bring to each collaboration. Medical practice needs more people like you who keep that spark in their eyes, no matter how gloomy the circumstances.

Marcus Ong. I'm so glad we ended up in that cab to the conference together. You are a powerhouse and your faith and commitment are admirable. When the going gets tough,... you reminded me how easy it can be to keep going. Quitting just doesn't seem like a real option anymore.

I wish to thank all of my colleagues at UT, old and new. Thank you Ada for all your support. Thanks to Michael, Julia, Nolen, Johnny, Koray, Miles, Yashar, Brandt, Peter-Paul, and especially Mieke.

Special thanks for Janjaap's emergency IT support that helped get the book published in time and to Guyaum for helping me get out of France in time.

Saskia Nagel. Thank you for believing I would do a good job on this project. Thanks especially for all the care and support you offered me and many others.

The same counts for Willem Lemmens. Thank you for your support and for being the first, many years ago, to ever suggest I get a PhD—even if, at the time, I laughed at the idea.

Ezio, I am so grateful for my time at UCPH. I loved the weekly philosophical exchange. So glad I got to meet Ji-Young, Andrea, and Rikke too. Thanks to Anja, Klaus, Henriette, Christoffer, Isaac, Anne, Signild, Jan Kyrre, and all of the colleagues at the Center for Medical Science and Technology Studies for their collegial spirit and valuable contributions.

Thanks to Bernike Pasveer, Govert Valkenburg, Harro van Lente and all the peers at WTMC with whom I shared a fantastic time and a wonderful education. The mere memory of that old monastery elicits a sense of serenity and reflection in me and our exchanges, both personal and professional, always sparked creativity and wonder. I loved our time there.

Thanks also to Sally Wyatt for brilliantly directing WTMC, with a lot of humor too, and for being such an admirable source of inspiration.

While I am bound to forget some important contributors, I am fully aware that many others have helped me along the way. At conferences and workshops, as well as through personal conversation... Heidrun Åm, Fern Wickson, Erik Fisher, Ulrike Felt, Joel Anderson, David Velleman, and David Enoch,... to name a few. Thank you.

To everyone at EPET—Lotte, Danielle, Merel, Mira, Karen, Ricky, Tamar, Nora, Darian, Darryl, Tsjalling, Ties,... and especially Simone. Thank you so much for the great gatherings and all your input, feedback and exchange.

Friends and colleagues, I've been lucky to not have to distinguish much between the two.

While I feel fortunate that I can consider most of my colleagues friends, I have many friends who are not colleagues, and yet they have contributed in their own ways. Gai, Betty, Tom & Zoë, Imoa, Koen, Sophie, Deepak & Sunaina, Deva, Martin, Dain, Daniel, Mar, Liat, Alex, Angelie & Timon, Liza, Jessica, Khalid, Lena, Jiri, Antonina, Marija, Fiona, Magali, Jorim, Vero & Alban, Mohamed, Ryan, Ricardo, Jeroen, Goedroen, Sanne, Hadewych, Hickxell, Mathilde, Yaniv, Pieter, Ravid, John, Conor, Courisse, Barbara & Sven, Julia Jay, and last but certainly not least, Tonia. All of you deserve your own dedication, whether you've offered me a place to stay, cooked for me, supported me through some rough patches, just showed that you cared,... most likely all of the above... I value each one of you not just for what you did for me but for your individual character, your values, and the way you stand for them. Having many friends, you get to be selective. If you're in my life today, you've earned your spot. Thank you for standing out, for being as remarkable as you are.

Brett and Aldona, my dear friends, I'm glad you got to meet each other. Your passing was heart-breaking and I miss you both. Brooklyn, New Haven, and Boston will never be the same again.

Chris, I have always appreciated your authenticity, your straightforwardness, your sense of humor, and your immense generosity. I am still not sure how I would've managed the move from Sweden to the Netherlands without you. Even more crucial, without you I wouldn't have been able to get out of France in time to print this book. Not to mention the times you drove hours to Utrecht to save me from myself or when you moved and stored all of my sister's stuff. 'Branding in de storm.' Thanks for being so unfalteringly trustworthy. Most people aren't so lucky to have a friend like you and I very much regret how little we've seen one another these last couple of years. I certainly hope we can soon make up for lost time.

Kirsten. While most people can count themselves lucky to have a friend like Chris... I am so lucky that I have two such friends. For over two decades now, I know I can count on you. Thank you for everything you've done for me, especially the support you offered during my PhD. You are one of the strongest, hardworking, and most resilient people I know and seeing you evolve through the years, even from a distance, has always been inspiring to me.

Alex. 18 years now... what can I even say. Parrots? Pirates? Thank you for showing up whenever it mattered and without being asked. I can't wait for the walking, and talking, and drinking coffee, and... Soon, with your kids. Later, when we're old, ... in so many different places around the world.

Gerard. Considering how much of the pressures you observed (and experienced) at the beginning of my PhD-track, I am glad to see you witness the end of it. Thanks for the years shared and for remaining in my life.

With all the history we share, I couldn't be happier that some of my lifelong-friends (and one very fresh one) have directly contributed to the making of this book.

Anya, how far we've come from Tel Aviv bars, exchanging surf and travel stories in places like New York and Hawaii we independently traveled and our shared adventures in Israel and new homes in Sweden and Italy and France and... what a life it has been. And while you were painting the artwork for this book, you were also creating life, growing David in your belly. At times, it feels like we gave birth at the same time. Thank you for being in my life.

Marco, we would have never found that type font without you so I guess you helped produce my albeit proverbial child after all. I am grateful for our history, nearly two decades of friendship, and all the work you put into each one of my enterprises over the years, at no cost. I think this particular endeavour will speak to you as you and I both know what it's like to lose a brother this way. I love you dearly and I thank you for everything that you are.

Henrik, the newest addition to the golden team. Thank you for having offered to read the penultimate draft before I even asked you to. Even more than the input you gave, I value your enthusiasm, not only for my work but for knowledge and learning in general. And for exercise and building stuff and humor and ice cream and sustainability and giving and discgolf and cooking and life. Please never stop.

Carro. Where do I even begin? How much I miss your cute presence in all the places that aren't Malmö. I am so glad that you visited me in Enschede and that my affiliation with the University of Copenhagen brought me so much closer to you. I often think of those few days where we got to spend time together and walk and talk, just the two of us. Thank you for your friendship, your humor, the conversations, the meditations we shared, and yes, ... the many hours of work you cheerfully put into turning google documents into a publishable book, well into the night. You must be some kind of superstar! You and Henrik have been a blessing in the final stages of my PhD. From the peace and nurturing stability you and your home offered and the meals you (mostly Henrik) prepared (effortlessly turning me back to a complete vegetarian, I might add!), to the very tangible efforts you made to get this book finalized... You have been home to me. I thank you from the bottom of my heart.

David. Merci pour l'amour et la lumière que tu m'as offert durant cette dernière année. J'ai aimé grandir à tes côtés. J'ai aimé te voir grandir aux miens. Peu importe ce qu'on deviendra, tu auras ta place dans mon cœur. Yalila, sache que de tes quatre à tes cinq ans, tu as été un rayon de soleil pour moi. Je suis curieuse de te voir grandir, toi aussi... Gracias, Sonia, por la magnanimidad con la que terminaste dándome mi espacio y tu reconocimiento. Preciosa!

As surprising as this may be to those who know me, I would also like to thank my father. Perhaps the greatest extender of all, I won't deny that there is also just plain *good* that I inherited from you. Without your influence, I likely wouldn't have cared as much for the artwork of this book, nor would I have had such a general strong passion for the aesthetic in art as well as in life. To an extent, your vices were period-specific. In some ways, you were just ahead of your time and you might have fared better today. I recognize myself in you in my adventurous lifestyle, my desire to excel at everything I do (and for that everything to include a lot of variety), and my unusual orientation towards romantic love and beauty. Finally, your death, and the process that preceded it at the end of the first year of my doctorate, was inconceivably valuable to my understanding of medical practice after cardiac arrest. My brother and sisters' experience of your demise forced me to step away from the perspective of the medical professionals at a time when the fieldwork was making me biased in their favor. The fact that you lost consciousness in a country where withdrawal of life-sustaining treatment was prohibited only reaffirmed my intuition

that cultural and contextual comparisons would be essential to this work. I have no regrets about how things ended between us but I also know that you'd be proud and happy to hold this book in your hands.

Even if these acknowledgments also reflect that the years spent working on a doctoral thesis were no bed of roses, no pleasure cruise - yes, I'm quoting Queen - I feel extremely blessed and fortunate for the opportunity and I thank you all.

Samenvatting:

Verantwoord Voorspellen bij Kritieke Onzekerheid

een epistemische analyse van neuroprognostische innovatie voor coma na hartstilstand

De middelen om een prognose te stellen bij patiënten in coma na hartstilstand zijn beperkt, en het resultaat is altijd onzeker. Die onzekerheid brengt een moreel dilemma voor familie en professionals met zich mee: als de behandeling op basis van de prognose wordt gestaakt, wordt mogelijk een leven van goede kwaliteit verhinderd, maar als er wordt doorbehandeld kan dat resulteren in een leven van zeer minimale kwaliteit. Innovatie van prognostiek wordt dan ook gedreven door de wens om de onzekerheid van de prognose, en daarmee het handelingsdilemma, te verminderen.

Dit proefschrift analyseert *of* en, zo ja, *hoe* nieuwe technologie een verantwoorde innovatie kan zijn van prognostiek voor patiënten in postanoxisch coma na hartstilstand. Specifiek heb ik gekeken naar de ontwikkeling van een nieuwe prognostische test op basis van continue electroencephalogram (cEEG) metingen.

Om te bepalen onder welke voorwaarden cEEG een verantwoorde innovatie van de prognostiek voor patiënten in coma na hartstilstand kan zijn, heb ik onderzocht welke zorgen betrokkenen (naasten, zorgprofessionals en ex-patiënten) in de huidige prognostische praktijk hebben. Ook heb ik verkend in hoeverre en hoe cEEG deze zorgen zou kunnen adresseren, en of er misschien ook nieuwe problemen kunnen ontstaan. Op basis van die verkenning heb ik aandachtspunten voor verantwoorde innovatie geformuleerd en suggesties gedaan voor hoe een verantwoorde ontwikkeling van cEEG te bevorderen. Veel van de aandachtspunten die ik heb geïdentificeerd zijn niet specifiek voor cEEG, maar ook relevant voor andere innovaties in de neuroprognostiek.

Voor dit alles heb ik een combinatie van verschillende methoden gebruikt. Naast onderzoek van de medische, filosofische en ethische literatuur heb ik uitgebreid etnografisch veldwerk verricht, op verschillende locaties en in verschillende landen. Dat veldwerk bestond uit observaties van de praktijk op de Intensive Care (IC) en in revalidatie- en zorgcentra, semi-gestructureerde interviews en enkele workshops met betrokkenen. Ook heb ik, samen met collega's, een kwantitatieve empirische studie gedaan door een survey te ontwikkelen die is uitgezet onder Nederlandse burgers. Het empirische werk heb ik steeds gecombineerd met conceptuele analyse. De wisselwerking tussen het empirische en het conceptuele hebben de analyse versterkt doordat de twee elkaars zwakke plekken aanvullen. Door de verscheidenheid in theoretische benaderingen en methoden hebben de hoofdstukken ook een verschillend karakter.

In hoofdstuk 1 plaats ik mijn onderzoek tegen de achtergrond van recente wetenschappelijke discussies over 'verantwoorde innovatie', oftewel 'responsible (research and) innovation' (RI). Ik motiveer mijn benadering van 'verantwoord innoveren' en leg uit waarom (en in welke mate) deze verschilt van de bestaande benaderingen. In die bestaande benaderingen keren drie veronderstellingen steeds terug: (1) Opkomende technologieën moeten worden beoordeeld vanwege hun radicale nieuwheid en onvoorspelbaarheid. (2) Een vroege beoordeling is noodzakelijk om het innovatietraject te beïnvloeden. (3) Anticiperen op het onbekende is nodig om je voor te bereiden op het onvoorspelbare. Ik beargumenteer dat deze drie aannames niet gelden voor wat ik 'liminale innovatiepraktijken' noem: innovaties die worden gedefinieerd door een continue heen en weer gaan tussen experiment en implementatie, tussen onderzoek en praktijk, en tussen eerdere en toekomstige innovaties. Allereerst (1) hebben technologieën die centraal staan in liminale innovatiepraktijken andere kenmerken dan die welke doorgaans worden toegeschreven aan opkomende technologieën. Bovendien (2) is het liminale innovatietraject anders, waardoor continue evaluatie en vormgeving ook lang na implementatie nog mogelijk is. Ten slotte (3) vereisen deze verschillen een heroriëntatie in RI-benadering. In plaats van anticipatie op het onbekende en onzekere dient verantwoorde liminale innovatie terug te keren naar observatie van het bekende en voorspelbare.

Hoofdstukken 2, 3 en 4 zijn dan ook gebaseerd op uitgebreid empirisch onderzoek van liminale innovatiepraktijken; prognostische praktijken op IC's waar cEEG-monitoring werd uitgevoerd in het kader van wetenschappelijk onderzoek. Ik heb op twee Nederlandse IC's observaties verricht en daar de zorg voor meer dan twintig patiënten in coma na hartstilstand kunnen volgen. Voor vier van die patiënten (met verschillende prognoses en verschillende uitkomsten) heb ik steeds vijf betrokkenen geïnterviewd, waaronder een tot drie naasten, twee of drie zorgverleners, en waar mogelijk de patiënt zelf. Bovendien heb ik ook observaties en interviews gedaan in een liminale innovatiepraktijk op een IC in de Verenigde Staten, en op verschillende locaties in Israël. Dit vergelijkende veldwerk hielp mij om de Nederlandse praktijk in perspectief te plaatsen.

In hoofdstuk 2 gebruik ik mijn observaties en interviews om het klinische dilemma dat ik in de inleiding al beschreef uitgebreider te analyseren. Ik beschrijf hoe de betrokkenen dit dilemma ervaren, en belicht welke zorgen en vragen het voor hen oproept. In de praktijk wordt prognostiek gevolgd door de keuze om de behandeling (a) te staken, (b) voort te zetten, of (c) die beslissing te vervroegen/uit te stellen. Die keuze blijkt te worden gebaseerd op een zorgvuldige afweging van mogelijke schade voor de patiënt, de familie van de patiënt en de samenleving als geheel. Mijn empirisch onderzoek toont echter aan dat al deze overwegingen niet alleen afhangen van de betrouwbaarheid van de voorspelde uitkomst, maar ook van de vraag of die voorspelde uitkomst de nadelen en voordelen die op het spel staan

correct weergeeft. Mijn analyse van de zorgen van de betrokkenen en hoe ze deze wegen in de besluitvorming, laat zien dat het vermijden van mogelijke schade voor een bepaalde patiënt antwoorden vereist op twee heel verschillende vragen: (1) welk resultaat kan worden bereikt en (2) zal dat resultaat acceptabel zijn? Pogingen om neuroprognostische onzekerheid te verminderen, zijn echter uitsluitend gericht op het beantwoorden van de eerste vraag.

In **hoofdstuk 3** link ik de twee vragen aan discussies over medische zinloosheid en ga ik dieper in op hoe betrokkenen in de klinische praktijk antwoord proberen te geven op de tweede vraag: welke uitkomst (welke kwaliteit van leven) acceptabel is. De beslissing of een levensverlengende behandeling van een patiënt in coma moet worden stopgezet, wordt meestal gemotiveerd door de overweging dat die behandeling 'medisch zinloos' is. In de medische en ethische literatuur over medische zinloosheid wordt onderscheid gemaakt tussen fysiologische en normatieve (of kwantitatieve versus kwalitatieve) zinloosheid. Onzekerheid met betrekking tot fysiologische zinloosheid wordt algemeen erkend en er zijn neuroprognostische tests ontwikkeld om die onzekerheid te verminderen. Mijn veldwerk laat echter zien dat oordelen over normatieve zinloosheid vaak net zo onzeker zijn: er is onzekerheid over hoe de uitkomst door de patiënt zal worden ervaren en geëvalueerd. Bovendien is het onzeker hoe omgevings-, contextuele en institutionele factoren die ervaring en evaluatie kunnen beïnvloeden. Ik beargumenteer daarom dat elke serieuze poging om neuroprognostische onzekerheid te verminderen duidelijk onderscheid moet maken tussen: (1) onzekerheid of de voorgestelde behandeling fysiologisch het gewenste effect kan bereiken (fysiologische onzekerheid) en (2) onzekerheid of de fysiologische effecten van de behandeling een acceptabele uitkomst zullen vormen (normatieve onzekerheid). Ik concludeer dat die laatste onzekerheid normatieve zinloosheidsoordelen zo onbepaald kan maken dat het nauwkeuriger is om te zeggen dat we normatieve overwegingen moeten voorspellen, in plaats van dat we ze kunnen identificeren. Mijn conclusie is dat de neuroprognostische praktijk dus niet één, maar twee voorspellingen vereist. Verschillende partijen kunnen ook verschillende expertise bieden om dergelijke uiteenlopende voorspellingen te informeren.

Zoals gezegd worden nieuwe prognostische tests zoals cEEG ontwikkeld met de ambitie het klinische dilemma bij besluitvorming over de behandeling van patiënten in coma na hartstilstand te verminderen. In **hoofdstuk 4** onderzoek ik hoe plausibel het is dat het gebruik van continue EEG-monitoring het gemakkelijker zal maken om te beslissen of een levensverlengende behandeling al dan niet moet worden stopgezet. Daarvoor breng ik in kaart welke effecten cEEG heeft als het onderdeel wordt van een sociaal-technisch netwerk op een afdeling intensive care. De analyse richt zich op (1) de manier waarop testresultaten worden geconstrueerd, (2) het

soort beslissingsondersteuning dat deze resultaten bieden, en (3) hoe cEEG de communicatie tussen professionals en familieleden beïnvloedt. Ik beargumenteer dat cEEG in sommige gevallen inderdaad onzekerheid kan wegnemen of verminderen, maar dat de technologie ook nieuwe zorgen oproept. Bovendien wijs ik erop dat de feitelijke effecten van cEEG in de praktijk afhangen van hoe cEEG in de praktijk wordt ontworpen en geïmplementeerd. Voor een verantwoorde implementatie zou bij de verdere ontwikkeling van cEEG bijvoorbeeld meer aandacht moeten gaan naar de betrouwbaarheid van de interpretaties van het EEG-signaal en naar de invloed van de user-interface op de interpretatie en communicatie van uitkomsten. Verder creëert cEEG ook een derde, 'grijze' uitkomstcategorie: EEG-patronen die noch duidelijk gunstig, noch duidelijk ongunstig zijn. Meer onderzoek naar de patiënten met deze uitkomst is gewenst. Een aantal adviezen zijn niet specifiek voor cEEG, maar relevant voor alle neuroprognostische innovatie. Zo adviseer ik bijvoorbeeld transparanter te zijn over het evaluatieve karakter van uitkomstmaten en de evaluatieve beoordeling van mogelijke uitkomsten zoveel mogelijk te scheiden van de fysiologische uitkomstmaten.

In **hoofdstuk 5** staat de vraag centraal wat de voorkeuren van Nederlandse burgers zijn ten aanzien van de omgang met fysiologische en normatieve onzekerheid bij neuroprognostiek. Welke prognostische informatie zouden zij willen ontvangen, hoe zouden besluiten over het al dan niet staken van levensverlengende behandeling genomen moeten worden, en wat vinden zij een acceptabele kwaliteit van leven na coma? Onder leiding van collega's van de afdeling Health Technology Assessment en in samenwerking met collega's van de Klinische Neurofysiologie is een enquête ontwikkeld die is uitgezet onder een representatieve steekproef van 500 Nederlandse volwassenen. De resultaten van dit onderzoek geven aan dat de huidige manier waarop prognostische informatie ingezet wordt om te voorspellen wat in fysiologisch en normatief opzicht zinloos is, niet volledig in lijn is met de voorkeuren van de respondenten. Wat betreft fysiologische zinloosheid acht de helft van de onderzoekspopulatie staken van behandeling bij patiënten in coma na een hartstilstand op basis van onzekere fysiologische prognostische informatie onaanvaardbaar. Met betrekking tot normatieve zinloosheid geeft meer dan de helft van de onderzoekspopulatie er de voorkeur aan in leven te blijven in toestanden die in de medische literatuur, medisch onderzoek en de klinische praktijk als een 'slechte' uitkomst worden beschouwd, en dus reden voor het staken van behandeling kunnen zijn. Dit geeft aan dat de manier waarop prognostische informatie op dit moment op Nederlandse ICs wordt ingezet zou moeten worden heroverwogen, willen we een verantwoord gebruik van nieuwe prognostische tests kunnen waarborgen.

In de laatste twee hoofdstukken focus ik op de onzekerheden die gepaard gaan met 'self-fulfilling prophecies', oftewel het fenomeen dat voorspellingen hun eigen

waarheid produceren. Het potentieel voor 'self-fulfilling prophecies' (SFPs) in de geneeskunde wordt algemeen erkend. Het is een belangrijk thema bij prognostiek op de intensive care, omdat beweringen over het effect van de behandeling of over de verwachte toestand van de patiënt een zelf-ervullend effect kunnen hebben. Als de prognose van een patiënt slecht is, leidt dat vaak tot de beslissing om de behandeling te staken. Dan is de uitkomst van de patiënt inderdaad slecht, want deze overlijdt. De eerdere voorspelling is dus uitgekomen, maar het is onzeker of dat hoe dan ook gebeurd zou zijn, of dat dit uitsluitend het gevolg is van het staken van de behandeling. In de medische en filosofische literatuur over SFPs richt de interesse zich tot nu toe op voorspellingen die het effect van de behandeling of de uitkomst van de patiënt veranderen.

In **hoofdstuk 6** presenteer ik een alternatieve analyse van de werking van SFPs in de geneeskunde. Centraal daarbij staat het argument dat SFPs altijd problematisch zijn, ongeacht of ze de uitkomst veranderen of niet. Omdat SFPs altijd waar zijn, wordt de kwaliteit van de voorspelling vaak niet eens in twijfel getrokken. Maar zelfs wanneer men het probeert kan men niet achteraf, contrafactisch, controleren in hoeverre de basis van de voorspelling correct was. Dit komt omdat we geen onderscheid kunnen maken tussen 'transformatieve SFPs' die de uitkomst veranderen en 'operatieve SFPs' die dat niet doen. Op basis van deze analyse maak ik twee punten. Allereerst produceren SFPs, omdat ze altijd hun eigen vervulling inhouden, geen betrouwbare feedback. Het niet kunnen vaststellen of de basis van de voorspelling juist of gebrekkig was, betekent dat men niet kan leren van mislukkingen, maar ook niet van 'gepercipieerd' succes. Dit is het centrale epistemische gebrek van SFPs. Ten tweede betoog ik dat dit gevolgen heeft voor de verantwoordelijkheid van onderzoekers en besluitvormers. Mijn punt is niet zozeer dat SFPs de ruimte laten aan voorspellers en besluitvormers om verantwoordelijkheid te ontlopen voor de gevolgen van hun voorspelling; er kunnen namelijk goede redenen zijn om de voorspelling alsnog als basis voor besluitvorming te gebruiken. Het is problematischer dat voorspellers en besluitvormers vanwege SFPs verantwoordelijkheid ontlopen voor hun falen om te leren.

In **hoofdstuk 7** onderzoek ik de implicaties van de analyse uit hoofdstuk 6 voor SFPs in de neuroprognostiek. Adequate evaluatie van neuroprognostische praktijken vereist dat we eerst alle SFPs goed identificeren, inclusief degenen die de uitkomst van de patiënt niet veranderen. Een dergelijke evaluatie houdt in dat men zich niet alleen meer richt op fout-positieve testresultaten, maar op de invloed van alle positieve testresultaten, juist of fout. Ten tweede moeten we binnen de neuroprognostiek twee verschillende SFPs identificeren; één door fysiologische onzekerheid en één vanwege normatieve onzekerheid. Een voorspelling dat de fysiologische uitkomst slecht zal zijn, maakt dat de uitkomst ook daadwerkelijk

slecht is, los van of dat anders ook zo zou zijn. Een normatieve voorspelling dat de fysiologische uitkomst slecht ervaren zal worden, vergroot de kans dat deze slecht ervaren zal worden. Omgekeerd zal een voorspelling dat de uitkomst als goed kan worden ervaren, de kans vergroten dat deze inderdaad als goed ervaren wordt. Ten slotte moeten we de effecten van fysiologische en normatieve SFPs in zowel acute als chronische zorg analyseren op twee onderling verbonden niveaus; ten eerste met betrekking tot de medische uitkomsten voor individuele patiënten, en ten tweede met betrekking tot de gevolgen voor alle leden van de patiëntengroep. Ik beargumenteer dat SFPs, afgezien van dat ze soms resulteren in ongewenste resultaten, het leren belemmeren. Dit maakt besluitvormers geneigd dezelfde fouten te herhalen, zowel met betrekking tot fysiologische als normatieve voorspellingen van medische zinloosheid. Geautomatiseerde voorspelling door middel van machine learning-technieken kan deze problemen verder verergeren. Ik laat zien hoe SFPs ten gevolge van fysiologische en normatieve voorspellingen de kwaliteit van medische praktijken in gevaar brengen, en innovatie en beleid op significante manieren compliceren. Ten slotte stel ik interventies voor om verkeerde prognoses te identificeren en de daaruit voortvloeiende feedback loops in onderzoek te doorbreken.

In de conclusie reflecteer ik op de bevindingen van dit onderzoek en hun bredere relevantie. Ik doe dit aan de hand van de onderzoeksvragen, door deze concreet te beantwoorden. Als motivatie voor mijn hoofdvraag 'op welke manier neuroprognostiek verantwoord te innoveren voor patiënten in postanoxisch coma na hartstilstand' schets ik eerst kort de problemen die zich in de praktijk voordoen. Omdat behandelingsbeslissingen op basis van prognostische informatie worden belemmerd door onzekerheid wordt er gewerkt aan het innoveren van neuroprognostiek. Om te kunnen analyseren welk effect prognose op basis van cEEG zou kunnen hebben op de huidige praktijk, geef ik een overzicht van de bestaande zorgen van betrokkenen in de zorg voor patiënten in postanoxisch coma in Nederland. Daaruit blijkt dat betrokkenen beslissingen op basis van onzekere voorspellingen over de neurologische uitkomst moeilijk vinden. Bovendien is er ook onzekerheid over hoe die uitkomst door de patiënt ervaren zal worden. Duidelijk communiceren blijkt in dit uitdagende proces essentieel. Na de analyse vat ik mijn adviezen samen voor hoe prognose op basis van cEEG best ontwikkeld en geïmplementeerd kan worden, zodat deze zoveel mogelijk rekening houdt met de bestaande zorgen en mogelijke onwenselijke effecten. Veel aandacht moet geschonken worden aan wie de cEEG-signalen interpreteert en hoe, en hoe deze interpretaties gecommuniceerd worden. Tenslotte weid ik uit over de bredere relevantie van mijn bevindingen. Eerst laat ik zien hoe mijn inzichten over liminale innovatiepraktijken niet enkel voor innovatie in klinische settings in het algemeen relevant blijken, maar ook voor innovaties in andere domeinen. Vervolgens laat ik zien waarom mijn analyse van self-fulfilling prophecies ook

relevant is voor praktische en geautomatiseerde voorspellingen in het algemeen. Een laatste overkoepelende opmerking gaat over het algemene belang van blijvende onzekerheid bij voorspellingen en hoe verschillende soorten onzekerheid daarbij een rol kunnen spelen. Ik concludeer dat afzien van voorspellingen over gezondheid en (risico op) ziekte soms een wijs alternatief kan bieden. Uiteindelijk duid ik aan welke van mijn adviezen niet enkel gelden voor het cEEG, maar voor neuroprognostische innovatie in het algemeen. Het gaat hierbij voornamelijk om epistemische kwesties. Allereerst is het noodzakelijk om fysiologische en normatieve onzekerheid in zowel onderzoek als zorg uit elkaar te trekken. Het is daarbij cruciaal uitkomsten niet langer in evaluatieve termen (bv. gunstig/ongunstig of goed/slecht) uit te drukken, maar zo veel mogelijk in beschrijvende termen (bv. CPC of GOS categorieën). Ten tweede moeten self-fulfilling prophecies in voorspellingen van zowel fysiologische als normatieve zinloosheid serieus aangepakt worden. Ik benadruk daarbij het belang van *leren* en *juist* leren als hoeksteen van verantwoorde innovatie.

"Deciding whether and when to use a new tool can be one of the most difficult challenges in clinical practice."

Kenneth W. Goodman – Ethics, Medicine, and Information Technology

"The only certainty, whether spoken or not, is that the doctors, nurses, and technicians are fighting not only death, but their own uncertainties as well. In most resuscitations, those uncertainties can be narrowed down to two main questions: Are we doing the right things? and, Should we be doing anything at all?"

Sherwin B. Nuland – How We Die