

**Paradigms in Operation: Pharmaceutical Benefit  
Assessments in England and Germany**

Katharina Kieslich

University College London (UCL)

Submitted for the Degree of Doctor of Philosophy in Political Science

## **Declaration**

I, Katharina Kieslich, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I can confirm that this has been indicated in the thesis.

Katharina Kieslich

## **Abstract**

The assessment of the benefits of pharmaceutical products through health technology assessments (HTAs) has become a feature of health care decision-making in numerous OECD countries, including England and Germany. Assessment outcomes vary between countries but, to date, there is a lack of research on the factors that affect those assessments. This thesis addresses this shortcoming by examining what determines the outcome of pharmaceutical benefit assessments in two countries that employ formalised HTA procedures. It takes a novel theoretical approach by employing a framework of policy paradigms to explain an empirical phenomenon other than policy change.

The study presents a qualitative analysis that compares the reasoning processes that led to assessment outcomes in ten of the same cases of pharmaceuticals in England and Germany. It finds that benefit assessment outcomes are determined by how a similar set of themes around evidence gets interpreted and framed by a HTA body, e.g. the National Institute for Health and Care Excellence (NICE) in England and the Federal Joint Committee (FJC) in Germany. The study explains the differences in addressing a similar set of themes around evidence by reference to different HTA paradigms that are applied, namely a cost effectiveness paradigm in England and a patient relevance paradigm in Germany.

The thesis makes a significant theoretical contribution because it demonstrates that policy paradigms can be used to explain empirical phenomena other than policy change. This requires an analysis of how paradigms are articulated in ‘normal decision-making’, much akin to Kuhn’s analysis on the connection between ‘normal science’ and paradigms. The study calls for an expansion of the current use of policy paradigms to include how they are operationalised in practice as this leads to a better understanding of the crucial elements of a paradigm.

## Table of Contents

Title Page	1
Declaration	2
Abstract	3
List of Tables	8
List of Figures	9
List of Abbreviations	10
Acknowledgements	12
<b>Chapter 1: Introduction</b>	<b>15</b>
1.0. Introduction	15
1.1. Health Care: Challenges and Reforms	20
1.1.1. Rationing and Priority Setting	23
1.2. Health Technology Assessment (HTA)	26
1.2.1. Controversies Arising from HTAs	28
1.3. Gaps in the Knowledge on HTA	32
1.4. Outline of the Study	33
1.5. Conclusion	39
<b>Chapter 2: Theoretical Framework</b>	<b>40</b>
2.0. Introduction	40
2.1. Policy Paradigms	43
2.2. Shortcomings in the Extant Literature	48
2.3. The Policy Paradigms of Pharmaceutical Benefit Assessments: A Conceptual Framework	55
2.4. Variables	58
2.5. Rules of Evidence	60
2.6. Alternative Theoretical Approaches	65
2.7. Conclusion	68
<b>Chapter 3: Research Design</b>	<b>70</b>
3.0. Research Aim	70
3.1. Methodology	71
3.2. Case Selection	75
3.3. Selection of Embedded Case Studies	80
3.4. Variables	81
3.4.1. Dependent Variable	84
3.4.2. Independent Variables	84
3.5. Data Sources, Collection and Analysis	87

3.5.1. Interview Process	89
3.5.2. Data Analysis	94
3.6. Concluding Remarks	96
<b>Chapter 4: Health Care, Pharmaceutical Policy and Pharmaceutical Benefit Assessment in England and Germany</b>	<b>98</b>
4.0. Introduction	98
4.1. Health Care in England	101
4.1.1. Financing	101
4.1.2. Principles, Structure and Health Care Provision	101
4.1.3. Cost Containment	103
4.1.4. Summary	103
4.2. Health Care in Germany	104
4.2.1. Financing	104
4.2.2. Principles, Structure and Health Care Provision	105
4.2.3. The Federal Joint Committee (FJC)	106
4.2.4. Cost Containment	107
4.2.5. Summary	108
4.3. Health Care in England and Germany: Concluding Remarks	109
4.4. Pharmaceutical Spending, Policy and Regulation	110
4.4.1. Pharmaceutical Licensing	111
4.4.2. Pricing and Reimbursement	112
4.5. ‘The Fourth Hurdle’: Health Technology Assessment (HTA)	114
4.6. The HTA Process in England and Germany	115
4.7. The HTA Process in England	118
4.7.1. The Dependent Variable: The Outcome of Pharmaceutical Benefit Assessments in England	121
4.8. The HTA Process in Germany	122
4.8.1. The Dependent Variable: The Outcome of Pharmaceutical Benefit Assessments in Germany	123
4.9. Institutional Features of NICE, the FJC and IQWiG: Comparative Summary	125
4.10. Conclusion	127
<b>Chapter 5: The Policy Paradigms of Pharmaceutical Benefit Assessments in England and Germany</b>	<b>128</b>
5.0. Introduction	128
5.1. Principles and Values within the wider Health Care Context	132
5.2. England: Principles and Values in Health Care	132
5.3. Germany: Principles and Values in Health Care	135
5.4. Comparative Views	138
5.5. Purpose and Goals of the HTA Processes	141
5.5.1. England and Germany: Purpose of HTA Processes	142
5.6. Principles, Criteria and Methods of HTA Decision-Making	144
5.7. England: Principles, Criteria and Methods of HTA Decision- Making	148
5.8. Germany: Principles, Criteria and Methods of HTA Decision-Making	151

5.9. The Case of Apixaban: How are Paradigms Operationalised?	156
5.10. Discussion	159
5.11. Conclusion	161
<b>Chapter 6: Policy Paradigms in Operation I: The Conceptualisation of Evidence</b>	<b>165</b>
6.0. Introduction	165
6.1. Empirical Results	167
6.2. Theme One: Permissibility, Quality and Validity of Evidence	171
6.2.1. Permissibility of other Types of Evidence	174
6.3. Theme One: The Cases of Cabazitaxel, Eribulin and Ipilimumab	178
6.3.1. Discussion	183
6.3.2. Research Findings	187
6.4. Theme Five: Suitability of Paradigms for ‘Special’ Cases	190
6.5. Theme Five: The Case of Retigabine	191
6.5.1. Research Findings	195
6.6. Conclusion	197
<b>Chapter 7: Policy Paradigms in Operation II: The Interpretation of Evidence</b>	<b>198</b>
7.0. Introduction	198
7.1. Theme Two: Choice of Comparator	199
7.2. Theme Two: The Case of Fingolimod	200
7.2.1. Discussion	203
7.2.2. Research Findings	207
7.3. Theme Two: The Case of Retigabine	208
7.3.1. Discussion	209
7.3.2. Research Findings	211
7.4. Theme Three: The Sub-Divisions of Patient Populations	212
7.5. Theme Three: The Case of Telaprevir	213
7.5.1. Discussion	215
7.5.2. Research Findings	221
7.6. Theme Four: Operationalisation of Thresholds and Algorithms under HTA Paradigms	223
7.6.1. Research Findings	226
7.7. Conclusion	227
<b>Chapter 8: Policy Paradigms in Operation III: Politics and the Articulation of Paradigms</b>	<b>229</b>
8.0. Introduction	229
8.1. Theme Six: Auxiliary Variables Emerging from the Empirical Evidence	230
8.2. Theme Six: Public Pressure as an Auxiliary Variable in England	231

8.2.1. Research Findings	238
8.3. Theme Six: Bargaining Powers and the Idea of Patient Relevance as Auxiliary Variables in Germany	240
8.3.1. Research Findings	246
<b>Chapter 9: Conclusion</b>	<b>250</b>
9.0. Introduction	250
9.1. Research Findings	254
9.1.1. Research Finding One: ‘Rules of Evidence’ Matter	254
9.1.2. Research Finding Two: HTA Paradigms Take Different Forms	255
9.1.3. Research Finding Three: A Similarity in Outcome Should not be Interpreted as a Similarity in Reasoning, or: Contrasting HTA Paradigms do not Necessarily Lead to Contrasting Outcomes	256
9.1.4. Research Finding Four: Core and Periphery Values Exist	257
9.1.5. Research Finding Five: Thresholds as the Expression of Paradigms in Normal Decision-Making	258
9.1.6. Research Finding Six: Paradigmatic Coherence Matters	259
9.1.7. Research Finding Seven: Ideational and Institutional Variables Matter	260
9.1.8. Research Finding Eight: Paradigms can Explain Empirical Phenomena other than Policy Change	261
9.2. Generalisability and Limitations of the Research Findings	262
9.3. Future Areas of Research	263
9.4. Policy Implications	265
9.5. Concluding Remarks	270
Appendix A: Overview of NICE and FJC Decisions	272
Appendix B: Interview Protocol	274
References	276
Additional Sources	295

## List of Tables

<b>Table 1.1.</b>	Summary of Research Findings	38
<b>Table 3.1.</b>	Overview of Pharmaceutical Products	81
<b>Table 3.2.</b>	Overview of Dependent and Independent Variables	83
<b>Table 3.3.</b>	List and Number of Stakeholders Interviewed	90
<b>Table 3.4.</b>	Number of Interviews According to Product and Country	91
<b>Table 3.5.</b>	Themes that Emerged from the Data Analysis	97
<b>Table 4.1.</b>	Health Care in England and Germany	100
<b>Table 4.2.</b>	Institutional Features of NICE, the FJC and IQWiG: Comparative Overview	117
<b>Table 5.1.</b>	The HTA Paradigms in England and Germany	130
<b>Table 5.2.</b>	Case Study: Apixaban	157
<b>Table 6.1.</b>	Themes Emerging from the Data (Detailed)	170
<b>Table 6.2.</b>	Permissibility of Evidence and the Hierarchy of Evidence	178
<b>Table 6.3.</b>	Case Study: Cabazitaxel	180
<b>Table 6.4.</b>	Case Study: Eribulin	181
<b>Table 6.5.</b>	Case Study: Ipilimumab	182
<b>Table 6.6.</b>	Case Study: Retigabine	192
<b>Table 7.1.</b>	Case Study: Fingolimod	201
<b>Table 7.2.</b>	Case Study: Telaprevir	214
<b>Table 8.1.</b>	Case Study: Abiraterone	233
<b>Table 9.1.</b>	HTA Paradigms in England and Germany	252
<b>Table 9.2.</b>	Paradigms in Normal HTA Decision-Making in England and Germany	253

## List of Figures

<b>Figure 1</b>	Total expenditure on health in OECD countries as % of GDP in 2011 (or nearest year)	21
<b>Figure 2</b>	Annual growth rate of total expenditure on health in real terms in the years 2000-2011	22
<b>Figure 3</b>	Total expenditure on pharmaceuticals and other medical non-durables as % of total expenditure on health in 2011 (or nearest year)	27
<b>Figure 4</b>	The Efficiency Frontier	147

## List of Abbreviations

ABPI	Association of the British Pharmaceutical Industry
AC	Appraisal Committee
ACD	Appraisal Consultation Document
AM-NutzenV	Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (The Ordinance for Assessing the Benefit of Pharmaceuticals)
AMG	Arzneimittelgesetz
AMNOG	Arzneimittelmarktneuordnungsgesetz (Act on the Reform of the Market for Medicinal Products)
BBC	British Broadcasting Corporation
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte (Federal Institute for Drugs and Medical Devices)
BMG	Bundesministerium für Gesundheit
BMJV	Bundesministerium der Justiz und für Verbraucherschutz
CCG	Clinical Commissioning Group
CDF	Cancer Drugs Fund
CP	Centralised Procedure
DCP	Decentralised Procedure
DH	Department of Health
DMP	Disease Management Programme
DRG	Diagnostic-Related Group
EC	European Commission
EMA	European Medicines Agency
ERG	Evidence Review Group
EU	European Union
F.A.Z.	Frankfurter Allgemeine Zeitung
FAD	Final Appraisal Determination
G-BA/FJC	Gemeinsamer Bundesausschuss/Federal Joint Committee

GDP	Gross Domestic Product
GP	General Practitioner
HAS	Haute Autorité de Santé
HRQoL	Health-Related Quality of Life
HTA	Health Technology Assessment
ICER	Incremental Cost Effectiveness Ratio
IQWiG	Institute for Quality and Efficiency in Health Care
MHRA	Medicines and Healthcare Products Regulatory Agency
MP	Member of Parliament
MS	Multiple Sclerosis
MTA	Multiple Technology Appraisal
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
OECD	Organisation for Economic Co-operation and Development
PAS	Patient Access Scheme
PCT	Primary Care Trust
PPRS	Pharmaceutical Price Regulation Scheme
QALYs	Quality-Adjusted Life Years
R&D	Research and Development
RCT	Randomised Controlled Trial
ROC	Return on Capital
SGB	Sozialgesetzbuch (Social Code Book)
SHI	Statutory Health Insurance
STA	Single Technology Appraisal
SVR	Sustained Virological Response
TA	Technology Appraisal
UK	United Kingdom
VBP	Value-Based Pricing
Vfa	Verband der forschenden Pharma-Unternehmen (German Association of Research-Based Pharmaceutical Companies)
WHO	World Health Organization

## Acknowledgements

My sincere gratitude, thanks and appreciation goes to my PhD supervisors, Professor Albert Weale and Professor David Coen. Throughout the past four years they listened to me articulate my ideas, they guided me with valuable input and they challenged me when my thoughts and analyses needed further development. In the many lengthy and intellectually stimulating supervisory meetings we had, Albert Weale and David Coen always encouraged me to do even better, to set my goals even higher and to challenge myself even further. The conversations and exchanges with my supervisors gave me the confidence that I needed to finish this thesis as well as the realisation that I should consider a potential career path that I had not considered as a serious option before. A special thank you goes to Albert Weale for involving me in so many of his projects, conferences and workshops and thereby introducing me to so many of my peers in the field. Above all, I thank Albert and David for being so supportive when I encountered some unexpected personal hurdles halfway through the process of completing my research. Their support and confidence in me helped me overcome these hurdles and for that I am very grateful.

I would like to thank the Konrad-Adenauer-Foundation in Germany for the PhD scholarship that supported me financially and academically for three years. Without this scholarship, this research would not have been possible. I also express my gratitude for the opportunity to participate in so many interesting seminars that the foundation offers for its scholarship recipients. The seminars enabled me to meet fellow scholars outside of my subject area and to engage in new and inspiring topics. During these seminars and throughout my funding period from the Konrad-Adenauer-Foundation I experienced what it means to be a part of a wider community of like-minded individuals. I found this experience very rewarding and I look forward to many more such exchanges within the foundation's community.

I would also like to express my gratitude to the University College London (UCL) and the Department of Political Science for the UCL Studentship for postgraduate work, which paid for my tuition fees through the department's IMPACT fund. Again, this research would not have been possible without the contribution from this studentship.

I extend my gratitude and appreciation to the individuals who participated in this research, especially to the interviewees in England and Germany who took time out of their busy workdays to talk to me about their experiences in the respective health technology assessment systems.

I would like to thank Allison Jones for proof reading my thesis just before submission. I know how much time this takes and I appreciate her help and comments very much.

A huge thank you goes to my best friend, Katharina B. McLarren. Her belief in me, along with her positive and happy views of life in general and the academic world in particular, provided me with so much personal, spiritual and academic support along the way. I am so grateful to Katharina for the many conversations about my PhD progress, for her encouraging words and for her ability to always make tough days easier and happy days even happier.

A special thanks for their support, love and laughter goes to two of my very best friends, Claire Yorke and Silvia Macri. Knowing that they are always there, that we can connect and communicate even when some miles are between us, has provided me with a huge sense of security. I feel incredibly lucky to count them as my friends and by being there, encouraging me and providing me with the much-appreciated distraction when needed, they contributed hugely to the support that I felt by so many people around me whilst finishing this thesis.

I would like to thank my friend and mentor, Dr. Rita Mohr-Lüllmann. Without her this research would not have been possible because it would have taken me much, much longer to discover that health policy is my chosen field of study and indeed for my career. Her professional insights on and her inspiring introduction to the world of health and pharmaceutical policy have shaped my career and academic progression more than I can say. Beyond that, Rita is a friend whose warmth and support has provided me with a great deal of strength and joy throughout the past years and for that, I am very grateful.

This thesis would not have been possible without the support, help and love from my family, old and new, near and far, extended and direct. Here, I especially thank Anni Kieslich, Heide and Bernhard Gimber, Christel Stegemann, Monika Iwersen, Dr. Susanne Zahn, Diane and Neil Seamark, Chris Seamark and Dr. Clare and Dr. Dave Seamark. In their unique and loving ways all these individuals have helped to keep me sane during what was a challenging and at the same time exciting and fulfilling time in

my academic career and personal life. I consider myself so very lucky to have had such a strong family support structure that I could rely on throughout the past four years.

I also thank my Dad, Paul, for encouraging me to do PhD as long as I can think. He was always of the opinion that I should do a PhD, that I am bound to do one and that there is no doubt about it. Even when I didn't know yet that this was the right choice, he did. His voice and belief in me accompanied me throughout the process of doing my PhD and it was especially strong in times of self-doubt. For that I am very grateful to my Dad.

Finally, my biggest gratitude, appreciation and love goes to my mum, Anette, and to my husband, Kevin. It is a gratitude that I cannot put into words. My mum and my husband were the ones that lived through the highs and lows of doing a PhD with me and they never lost their patience with me even when much of what I was saying at various points was nothing new to them. They discussed my research with me, they read my work and they relentlessly pushed me to believe in myself. They also provided me with the necessary and fun 'view outside the PhD world' during holidays and weekends. Most importantly though they always make me feel loved. For this, and for many, many other reasons that cannot all be mentioned here, I dedicate this thesis to my mother, Anette, and to my husband, Kevin.

# Chapter 1

## Introduction

### 1.0. Introduction

In October 2011 the National Institute for Health and Care Excellence (NICE) in England and the Institute for Quality and Efficiency in Health Care (IQWiG) in Germany published their assessment of a medicine called Ticagrelor<sup>1</sup> for acute coronary syndromes. The assessment was based on an evaluation of the scientific evidence in the form of data from randomised controlled trials (RCTs) and other studies submitted by the manufacturer of Ticagrelor. The pharmaceutical manufacturer submitted the same evidence in England and Germany (NICE, 2011; IQWiG, 2011). Based on the assessment of the evidence, NICE recommended the use of Ticagrelor for all of the four types of acute coronary syndromes that are distinguished for clinical purposes (NICE, 2011). NICE was satisfied that the expected health benefits for the eligible patient groups were sufficiently large to justify the costs incurred by the use of Ticagrelor. As a result, the costs for the use of Ticagrelor are covered by the National Health Service (NHS) in England. By contrast, IQWiG in Germany concluded there was a significant added benefit for only *one* of the four types of acute coronary syndromes (IQWiG, 2011). The benefit of Ticagrelor was assessed differently by NICE and IQWiG. How can this different assessment of the same pharmaceutical product be explained and why does it matter?

This thesis addresses the research question of what determines the outcome of pharmaceutical benefit assessments in health care systems that employ formalised health technology assessment (HTA) procedures. In doing so, it addresses the issue of how differences and similarities in assessment outcomes can be explained. In the context of health care policy this is an important issue because it gives rise to an empirical puzzle that is contrary to what one might reasonably expect if issue characteristics (Lowi, 1964) alone determined the outcome of public policy decisions. That is to say that one might expect that the benefits of a pharmaceutical product that has the same biochemical ingredients and characteristics worldwide would be assessed

---

<sup>1</sup> In line with the HTA organisations' use of the names of pharmaceutical products, the generic rather than the brand names of pharmaceutical products are used throughout this thesis.

similarly everywhere. However, comparisons of the outcomes of pharmaceutical benefit assessments between HTA organisations – organisations that review the evidence on the clinical and cost effectiveness of new medicines and other health technologies - have demonstrated that this is not the case (e.g. Kanavos, et al., 2010). While some benefit assessments of pharmaceuticals are similar, others are divergent. In exploring what determines similarities and differences in outcomes, this thesis seeks to offer novel empirical and theoretical insights into how evidential questions on new pharmaceutical products are framed in the context of different HTA paradigms, which ultimately provide a more in-depth understanding of the issues that determine the outcome of pharmaceutical benefit assessments.

For the policy process and the actors that engage in it, the question of how differences in benefit assessments of the same pharmaceutical products can be explained matters because different outcomes of benefit assessments lead to divergences in health care provision in countries that employ HTA procedures. For example, a negative assessment by NICE in England means that the NHS is not obliged to ensure access to the medicine or treatment in question. This implies that negative pharmaceutical benefit assessments can lead to a situation in which patients are denied access to a medicine or treatment. This in turn raises issues of political salience and ethical sensitivity, especially if access to the medicines in question is ensured in other countries or regions due to a different assessment of the benefit of the medicines.

The political salience and ethical sensitivity of the issues arising in the context of HTAs are exemplified by NICE's assessment of a new breast cancer drug called Kadcyła. In April 2014 NICE – the organisation that reviews evidence on the clinical and cost effectiveness of new medicines in England<sup>2</sup> – published a draft guidance document on Kadcyła. NICE concluded that the “breast cancer drug costing tens of thousands of pounds more than other treatments [is] ‘unaffordable’ for [the] NHS” (NICE, 2014). According to NICE Chief Executive, Sir Andrew Dillon:

[...] the reality is that given its price and what it offers to patients, it will displace more health benefit which the NHS could achieve in other ways, than it will offer to patients with breast cancer (NICE, 2014).

---

<sup>2</sup> NICE guidance usually applies to the NHS in England, but in selected circumstances it is also applied in Northern Ireland, Scotland and Wales (NICE, 2013).

Following NICE's publication of the draft guidance, a consultant oncologist described it as a "huge blow" (BBC, 2014) as "Kadcyla represents a significant advance in [...] breast cancer [...]" (BBC, 2014), whilst the representative of a breast cancer patient charity organisation said that "[...] we are concerned by the increasing number of people we support telling us how anxious they are about being able to access treatments when they need them" (BBC, 2014).

The principle that patients should have access to the medical treatments they need, when they need it, regardless of their socio-economic and financial background, is arguably one of the central pillars that modern health care systems are built on. However, the example of Kadcyla demonstrates that, in times of growing financial pressures on public budgets, it is proving increasingly challenging to ensure the timely access to treatment, including access to medicines, based on the criterion of medical need alone. This gives rise to a situation in which politically salient and ethically sensitive decisions on the allocation of health care resources have to be made. According to Freeman: "To guarantee access to health care means ensuring the availability of medicines, and doing so means addressing familiar distributional issues of who gets what, when, how" (Freeman, 2009, p. 244).

In order to address such distributional issues of health care access, in health policy a growing emphasis is placed on the assessment of the added value of medicines, both in medical and financial terms. Put differently, in an effort to find decision-making tools for making distributional choices, health policy-makers and decision-makers<sup>3</sup> have introduced requirements to evaluate the scientific evidence on a health care intervention or pharmaceutical product in order to decide whether access is justified in relation to the benefits and/or costs incurred. In a number of member states of the Organisation for Economic Co-operation and Development (OECD) this has led to the establishment of institutions – so-called health technology assessment (HTA) institutions – that are

---

<sup>3</sup> For the purpose of this thesis, health policy-makers are political representatives of government and opposition parties who are engaged in legislative activities that provide the statutory framework of health care. Health decision-makers include executive bureaucrats as well as actors at the operational level such as health care providers. Whilst assessments of medical interventions and pharmaceutical products emerge from statutory mandates in the majority of countries, the distinction between health policy-makers and health care decision-makers is made to highlight that such mandates can also arise from requirements of professional bodies such as physician associations or payers' organisations such as sickness insurance funds. The distinction serves the purpose of differentiating between the actors who determine the general direction of health policy and the actors who are in charge of making decisions within a given health care system on a daily basis.

commissioned with carrying out evaluations of the scientific evidence of interventions and pharmaceutical products in health care.

The examples of Ticagrelor and Kadcyra highlight that negative HTA evaluations of new pharmaceutical products may result in a certain product not being covered under a national health care scheme, which effectively means that patients may be denied access to certain medicines or other treatments. Denying patients access to health care treatment can have serious political consequences, especially when it is portrayed as a matter of injustice in the media and other public domains. For example, in 2008 patient campaigners protested against NICE's decision not to recommend a number of kidney cancer medicines on the grounds that they were not cost effective by saying that kidney cancer patients "deserved the right to life" (Walker and Batty, 2008). The protest against NICE was framed around questions of ethics, equity and human rights, essentially arguing that NICE was denying kidney cancer patients the right to life. This underlines the political salience and ethical sensitivity when it comes to deciding "[...] who gets what, when, how" (Freeman, 2009, p. 244). It also underlines the importance of understanding how outcomes of pharmaceutical benefit assessments can be explained in an effort to evaluate whether policy goals are being met and whether public criticism is justified.

Despite the political salience of the issue, empirical studies of HTA processes and outcomes remain incomplete. To date, studies on the subject predominantly focus on institutional structures and formal decision-making criteria. What is missing from these studies is an empirical account of the factors that determine the outcome of pharmaceutical benefit assessments<sup>4</sup>. For example, whilst criteria such as clinical and cost effectiveness are considered important to HTA decision-making processes, there is little understanding of how these criteria are prioritised, operationalised and balanced against other potential factors such as stakeholder views.

Moreover, an account of how questions of 'evidence' are formulated and interpreted in different national HTA contexts is missing from the extant literature.

---

<sup>4</sup> In this study the term 'health technology assessment' (HTA) refers to HTAs that are conducted on pharmaceutical products. In this sense 'HTAs' and 'pharmaceutical benefit assessments' are used interchangeably as the latter represents a form of HTA. The definition of HTA provided in this chapter and in chapter 4 shows that HTAs pertain to a wide array of health care interventions including medical devices and pharmaceutical products. However, the focus of this thesis is on the HTAs of pharmaceutical products and therefore both terms are employed to describe the formal evaluation of pharmaceutical products in countries that employ HTA procedures.

Given the centrality of ‘evidence’ to any HTA process, this is a surprising gap in the current knowledge on what determines the outcome of pharmaceutical benefit assessments. Additionally, there are, to the best of my knowledge, no comparative studies that trace how decisions on the same pharmaceutical products are arrived at in different countries. Given the good availability of data due to publicly available HTA decision documents and given the ability to control for issue characteristics (Lowi, 1964) such as different disease types, this is an area that this thesis seeks to address.

The thesis addresses the outlined gaps by offering an account of the factors that determine pharmaceutical benefit assessment outcomes based on the qualitative analysis of the decision-making process of ten of the same pharmaceutical products<sup>5</sup> that were appraised in England and Germany<sup>6</sup> in 2011 and 2012. Conducting a cross-national comparison on this issue offers a deeper understanding of the role that context plays to the outcome of what has been referred to as a case of “scientific-bureaucratic medicine” (Harrison, Moran and Wood, 2002).

The thesis seeks to contribute to the empirical knowledge of health policy generally, and HTA policy specifically, by providing an analysis of the factors that determine pharmaceutical benefit assessment outcomes. The analysis concludes that how concepts, ideas and criteria of HTA decision-making are interpreted in different contexts determines the outcome of pharmaceutical benefit assessments. The research findings illustrate that pharmaceutical benefit assessments are not technocratic, value-neutral processes that invariably result in similar appraisals of available evidence. Instead, they exemplify complex decision-making procedures in the realms of public policy that involve applying careful judgements to science on the basis of medical, scientific, economic, political and ethical considerations. What these considerations are, and how they are transformed into meaningful judgements around evidential questions, contributes to an understanding of what determines pharmaceutical benefit assessments in different countries.

The thesis’ theoretical contribution lies in the fact that it offers a new model for understanding HTA processes and their outcomes. The model is derived from an extension of theories on policy paradigms to the field of HTA. It demonstrates that

---

<sup>5</sup> See chapter 2 for an overview of the cases of pharmaceutical benefit assessments that were analysed.

<sup>6</sup> See chapter 3 for the description of the case selection process that led to the choice of England and Germany as case studies.

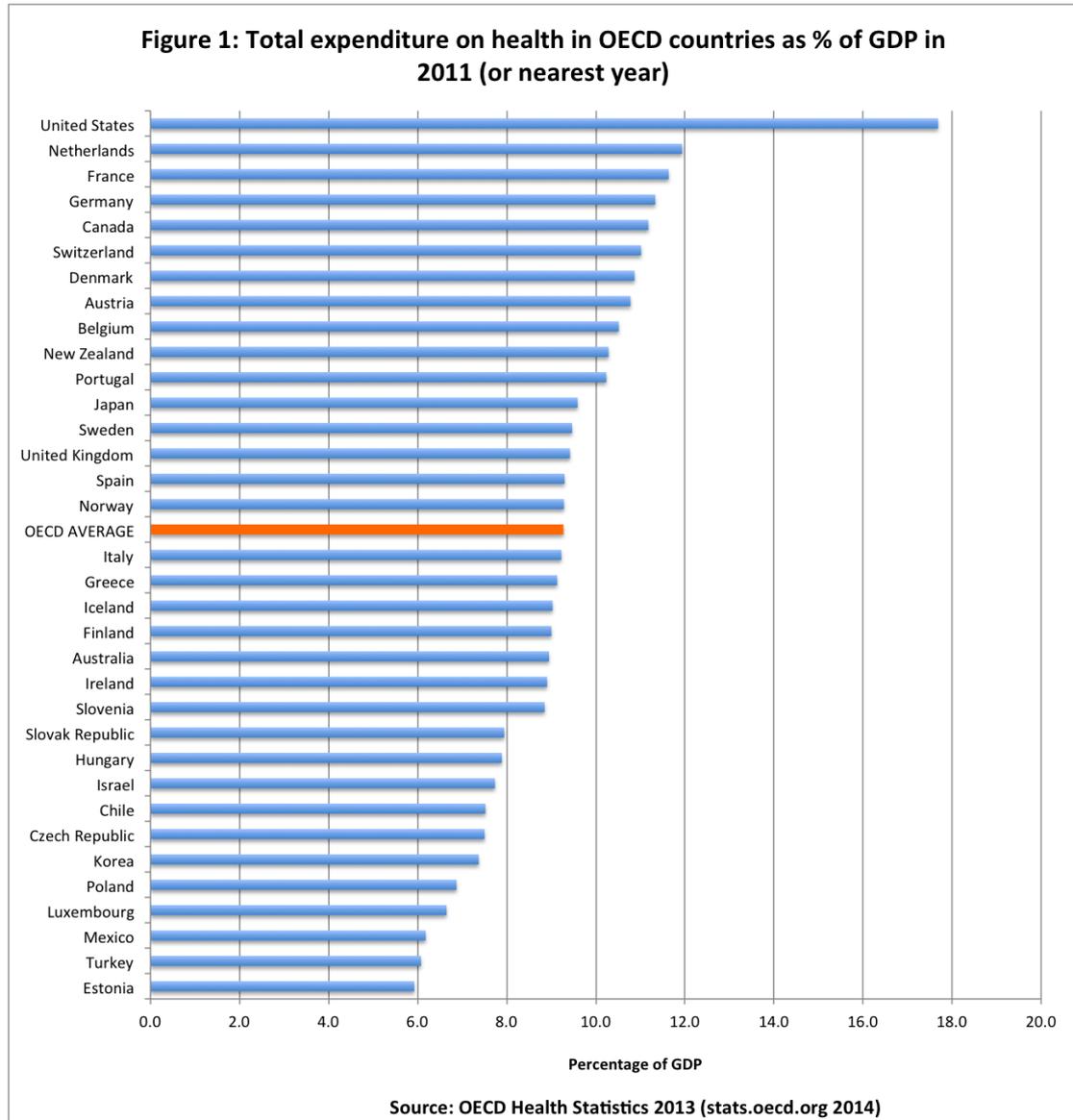
policy paradigms, conceptualised as intellectual constructs that determine what is or is not considered important in a HTA process, help explain both convergence and divergence of HTA outcomes. The theoretical approach is novel in that it applies theories that are commonly used to explain policy change to an empirical phenomenon other than policy change. In doing so, I demonstrate that the use of theories on policy paradigms can be expanded to explain a larger number of empirical questions than is currently the case. The adoption of policy paradigm frameworks resulted in the finding that a) the goals, values and ideas that shape an HTA system differ in different institutional contexts and that b) even ostensibly neutral concepts such as clinical effectiveness have divergent meanings in different contexts. This in turn impacts on the final outcome of pharmaceutical benefit assessments and can help explain why decisions on the same pharmaceutical product may vary between HTA agencies.

In the following sections I provide an account of the current challenges faced by health policy-makers and decision-makers. This provides the backdrop against which HTA processes to inform decision-making in health care have to be understood. Following this account, the chapter offers an explanation of HTA processes and their discussion in the extant literature. The research question that guides this study – that is, what determines the outcome of pharmaceutical benefit assessments in health care systems that employ formalised HTA procedures – emerges from this discussion. Finally, this introductory chapter provides the structural outline of this thesis and its main research findings.

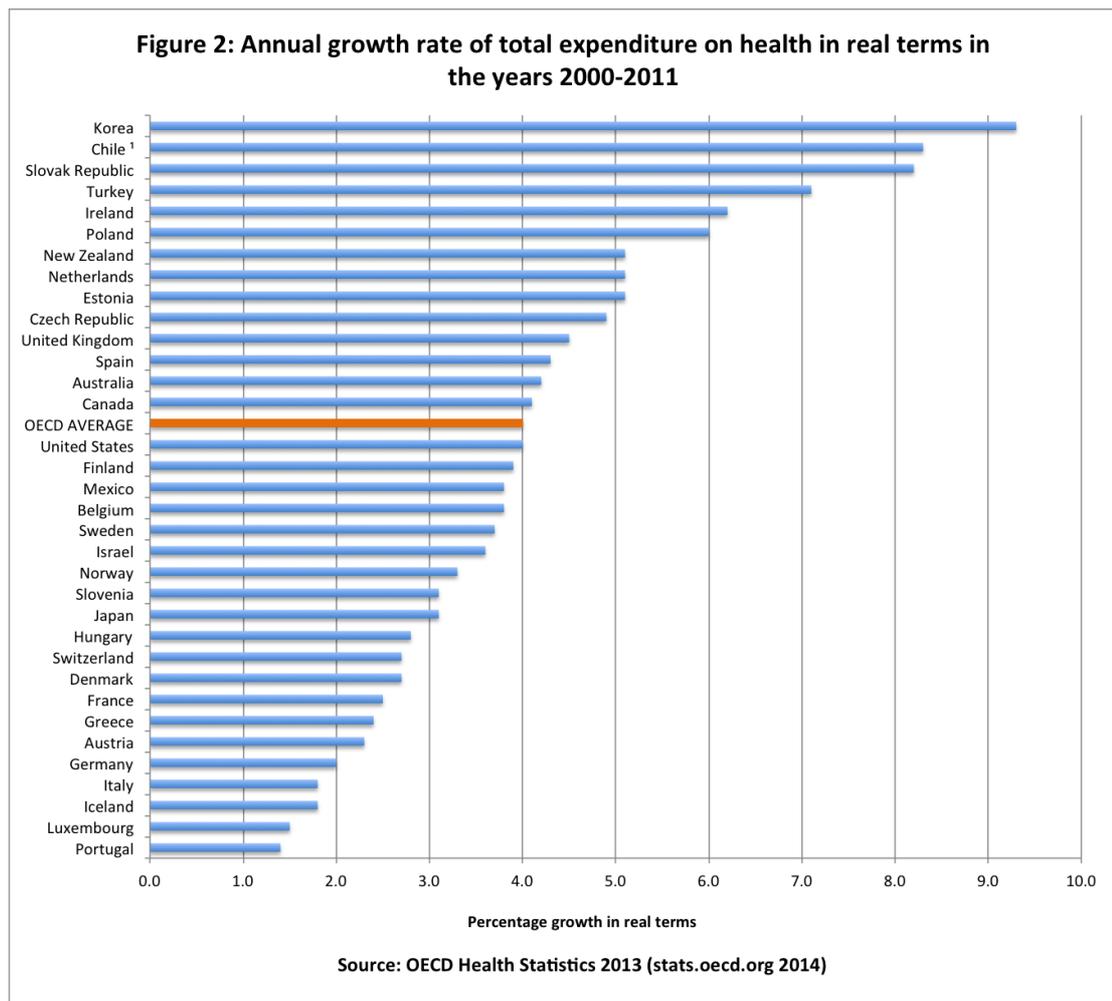
### **1.1. Health Care: Challenges and Reforms**

According to the extant literature the biggest challenge facing the health care state (Moran, 1995; 1999) emanates from financial pressures connected to rising expenditures. Access to pharmaceutical products is not the only matter of expense that public or private payers of health care need to cover. Other expenses arise such as the costs for the provision of hospital services as well as primary and community care services. The provision of such services incurs large health care expenditures in OECD countries (see figure 1). Moreover, the expenditures are increasing at a pace that exceeds that of inflation in a number of countries. According to the OECD (OECD, 2014), the annual growth rate of expenditure on health care in real terms between the

years of 2000-2011 was 4.1%, with the United States of America (USA) averaging at 4.8% and the United Kingdom at 4.5% (see figure 2).



The reasons for the continued rise in health care expenditure are multiple and emerge from various developments. Hanisch and Kanavos cite “[...] an ageing population, increasing incidence of chronic disease, persistent inequalities, and rising citizen expectations [...]” (Hanisch and Kanavos, 2008, p. 1) as factors that contribute to the financial strain that some health care systems find themselves in. Blank and Burau add the development of “[...] new medical technologies [...]” (Blank and Burau 2006, p. 265) to the matrix of key drivers for the rise in health care costs. Similar assessments can be found in Dixon and Poteliakhoff (2012), Freeman and Moran (2000), Ham and Honigsbaum (1998) and Ham (1997). The advancement of modern



medical technology, along with what Abraham refers to as the “expert patient or informed patient discourse” (Abraham, 2009, p. 935), whereby information on new medicines and therapies is readily available on the internet, leads to rising expectations of patients which further exacerbates the pressures on health systems to finance all that is medically possible.

Moran (1995; 1999) explains the challenges facing health care systems in capitalist economies as tensions arising from the phenomenon he calls the ‘health care state’. Moran (1995; 1999) argues that health care is more than an element of a state’s welfare activities in advanced industrial economies. In health care, the state assumes a central role, not just as the driver of welfare activities, but as an actor in the areas of industrial activity and distributional tasks. Moran (1995; 1999) refers to these state activities as “the three faces of the health care state”. In addition to shaping and guaranteeing the provision of health care to its citizens, states are concerned with supporting industrial activity, for example in pharmaceutical research and development, whilst also regulating the activities of, for example, health care professionals. It implies

that states face a number of competing demands and challenges that need to be addressed without jeopardising the relative standing of any one of them. Taken together, these tensions within the health care state appear to be facilitators of rising health costs.

The continuing trend of growing health care expenditure has led to cost containment efforts in the majority of OECD countries. In an effort to contain costs policy-makers have sought to uncover untapped efficiency savings by evaluating the health benefits and cost effectiveness of clinical services. The idea behind this is that such evaluations, or assessments, will shed light on what works best and what does not, in turn allowing for an informed decision on which services and medications to fund. In this environment, HTAs have become a synonym for evaluations to determine what works in health care and whether the costs of ‘what works’ are justifiable in relation to health benefits that can reasonably be expected. HTAs are evaluations of scientific evidence on a pharmaceutical product or other health care intervention with the aim to appraise it with regards to its therapeutic benefits and/or costs (e.g. Sloan, 1995; Nord, 1999). Be it in the form of institutions such as NICE in England and Wales or academic advisory groups based at universities, a number of health care systems now incorporate HTAs as a feature in their decision-making (Sorenson, 2009).

The introduction of HTA policies is part of a wider trend in health care, namely the trend towards rationing and prioritisation in times of growing financial pressures. The next section describes this trend in more detail in order to highlight a number of politically salient and ethically sensitive challenges that arise from it.

### **1.1.1. Rationing and Priority Setting**

According to Ham (1997) the move towards rationing and making difficult priority setting choices is characteristic of a third phase of health care reforms in OECD countries. It follows the first and second phases, which marked an effort to achieve cost containment at a macro level (1970s until the early 1980s) and measures to increase efficiency at a micro level (1980s until the 1990s) respectively. The introduction of HTA procedures thus has to be viewed within the context of a general trend in health care reform that focuses on prioritising or rationing services based on their clinical and cost effectiveness to address the continued financial pressures faced by health care systems. The need for setting priorities arises “because no country can afford to provide all its residents with every possible medical or public health intervention [...]”

(Littlejohns, et al., 2012, p. 286). Policy-makers are looking for ways to determine priorities in a manner that will ensure the efficient and fair use of resources. To put it differently: “Priority setting in health aims to determine what, in the context of limited resources, is most important” (Clark and Weale, 2012, p. 293).

How rationing and priority setting relate to one another is a matter of contestation. For example, Klein (2010) criticises the view that they are conceptually close because it leads to a situation in which the terms are used interchangeably in the literature. He argues for a clear distinction between the two concepts as:

Priority setting describes decisions about the allocation of resources between the competing claims of different services, different patient groups or different elements of care. Rationing, in turn, describes the effect of those decisions on individual patients, that is, the extent to which patients receive less than the best possible treatment as a result (Klein, 2010, p. 389).

Klein (2010) also distinguishes between explicit and implicit forms of rationing. Explicit forms of rationing include mechanisms such as waiting lists or the denial of medicines and delaying of treatments while implicit forms are harder to discern (Klein, 2010). This is because they might include mechanisms such as understaffing in hospitals or holding off on specialist referrals in order to keep expenditure in check (Klein, 2010).

In contrast to Klein, Coulter and Ham do not believe “[...] in drawing hard and fast distinctions between rationing and priority setting” (Coulter and Ham, 2000, pp. 1-2) as rationing “[...] has come to be employed to describe the variety of ways in which choices in health care are made whether they affect individuals, communities or countries” (Coulter and Ham, 2000, pp. 1-2). For the purpose of this research, Coulter’s and Ham’s approach is followed in that the distinction between priority setting and rationing is not operationalised in a strict manner, but is understood to be fluid. This is because the development of HTA policies is part of a wider policy-making discourse on rationing and prioritisation in which the boundaries between the two are not easily demarcated. That is to say that the outcome of HTAs may, depending on the observer’s viewpoint, be understood as either a form of rationing in that it denies patients access to a certain medicine if the appraisal of the medicine is negative, or it may be perceived as a form of setting priorities by prioritising the use of medicines that are deemed to meet the criteria that a HTA system has set for positive recommendations.

According to Coulter and Ham, rationing is “contentious” (Coulter and Ham, 2000, p. 11). It is contentious because setting priorities implies that one issue, one treatment, one health care area is given more attention than another which in turn suggests that not everybody’s expectations and needs can be met. Patients might be denied treatments or medicines. Littlejohns, et al. suggest: “In this context, priority setting should aim to produce allocation of healthcare resources that can be ethically justified, especially to those who lose out as a result of resource allocation decisions” (Littlejohns, et al., 2012, p. 286). As the controversies arising from priority setting have become clearer, the current literature notes a trend towards explicit prioritisation that focuses on “[...] making transparent the rationale for these priorities and basing resource allocation decisions on agreed-upon priorities” (Kenny and Joffres, 2008, p. 147).

One element of the trend towards explicit prioritisation is the introduction of so-called health technology assessments (HTAs) as an instrument to determine priorities and define the inclusion of services in health care benefit baskets. Some argue that HTA is considered a technique for determining priorities (Coulter and Ham, 2000, p. 11) while others label it a “[...] a new policy analytical tool for the health policy area” (Lehoux and Blume, 2000, p. 1085). As such HTAs need to be understood in the context of the wider debate on rationing and prioritisation in the current literature whilst giving rise to their own set of controversies that underline the view that a debate exists on “how it [priority setting] should be done” (Coulter and Ham, 2000, p. 11).

In the next section I provide a definition of HTA and outline some of the controversies that the process of HTA gives rise to. In doing so, I explain the contextual backdrop from which my research question - what determines the outcome of pharmaceutical benefit assessments in countries that employ formalised HTA procedures? – emerges.

## 1.2. Health Technology Assessment (HTA)

HTA procedures have become an important element of pharmaceutical reform efforts in OECD countries (e.g. Mossialos, Mrazek and Walley, 2004; Abraham, 2009; Sorenson, 2009; Maynard and Bloor, 2003). HTA is defined as:

The systematic evaluation of the properties and effects of a health technology, addressing the direct and intended effects of this technology, as well as its indirect and unintended consequences, and aimed mainly at informing decision making regarding health technologies (HTA Glossary, 2014).

Health technologies include “diagnostic and treatment methods, medical equipment, pharmaceuticals [...]” (EUnetHTA, 2014). HTAs are carried out on the basis of scientific evidence such as randomised controlled trials (RCTs) that provide an insight into the “properties and effects of a health technology” (HTA Glossary, 2014). The underlying idea is that the evaluation of available evidence on a given technology or pharmaceutical product will help determine which priorities to set and what medicines, treatments and surgical procedures to provide. Kanavos, et al. summarise this idea in the following way:

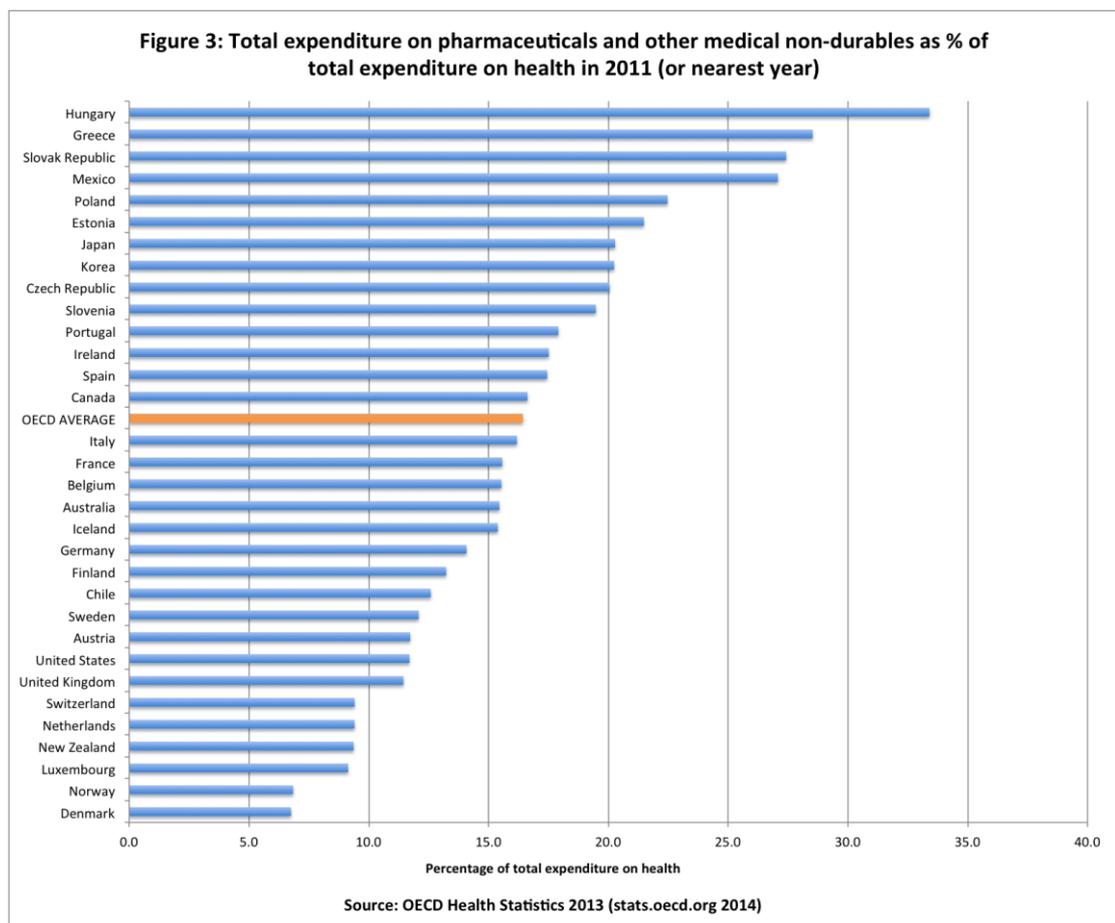
In an environment where resources are scarce, HTA agencies’ objective is to ensure access to safe and effective medicines, while managing health care expenditure in an efficient way by reimbursing clinically cost-effective treatments. In this discourse, pharmaceutical products are the main – but by no means the only – subjects of such appraisals (Kanavos, et al., 2010, p. 1).

The latter part of Kanavos, et al.’s conceptualisation of HTA’s objectives alludes to fact that while HTA can be used to inform decisions in different areas, there is a trend towards basing pharmaceutical coverage decisions on HTAs. Sorenson makes a similar point by arguing:

The use of HTA in pharmaceutical coverage decisions has grown substantially since the late 1990s and is likely to expand further, as national policy makers continue to face cost pressures and attempt to use evidence-based approaches to ensure effectiveness, efficiency and sustainability of their health systems (Sorenson, 2009, p. 4).

The trend towards the use of HTA to inform pharmaceutical coverage decisions forms the basis of the semantic use of HTA in this thesis. While I recognise that HTA can have a wider remit than pharmaceutical benefit assessments, in this thesis HTA refers to HTA procedures used to inform decision-making on pharmaceutical access or price setting. HTA is used interchangeably with pharmaceutical benefit assessments. That is to say that when the term ‘HTA’ is used I mean HTAs on pharmaceutical products, thus connecting the use of the term to the empirical focus of my research.

As highlighted by Kanavos, et al. (2010) and Sorenson (2009) pharmaceutical spending has been a prominent target of health care reforms generally and of rationing and prioritisation efforts specifically. There are three reasons for this. Firstly, pharmaceutical spending constitutes a large part of overall health care spending (see figure 3). The OECD average of spending on pharmaceuticals as a percentage of overall expenditure on health constituted 15.9% in 2012 (OECD, 2014). Secondly, the policy instruments for curtailing pharmaceutical spending are comparatively varied and provide policy-makers with a set of policy choices. For example, Freeman (2009) highlights patient co-payments, reference pricing, profit controls for the pharmaceutical industry and generic drug prescription as possible instruments to address rising pharmaceutical expenditures.



Thirdly, the nature of the pharmaceutical market gives rise to a distinct set of complex and competing considerations with which policy-makers need to engage (e.g. Maynard and Bloor, 2003; Mossialos, Walley and Mrazek, 2004). This includes balancing issues such as the need to ensure safe and affordable access to effective and innovative medicines whilst avoiding waste. It also includes supporting the pharmaceutical industry as a major provider of employment and an investor in research. The issues that this difficult balancing act gives rise to are reflective of the three faces of the health care state outlined by Moran (1995; 1999).

Evaluating the health benefits and cost effectiveness of pharmaceutical products has also become popular because of the fast pace of technological innovation in the area. The intellectual premise of assessing pharmaceuticals is a) that a new medicine does not, as a matter of course, represent a technological innovation unless its health benefits in comparison to current treatments are substantial and b) that even if a new medicine is innovative in its effect, the extent of the effect will determine whether its cost is justified. Such is the intellectual starting point of pharmaceutical benefit assessments, but as the following sections show, their use as a priority setting instrument is by no means uncontroversial.

### **1.2.1. Controversies Arising from HTAs**

The extant literature depicts several areas of controversy around HTAs, which can be divided into political/economic, methodological and societal/ethical dimensions. They centre primarily on the conceptualisation of HTA as a scientifically objective and value-neutral policy tool and on the fact that complex judgements still to be made by decision-makers in the process.

The political and economic dimensions of HTA arise from the high hopes that policy-makers have placed on HTA policies as remedies to some of the challenges facing health care states. Rationing and priority setting give rise to emotional debates, conflicts and protests (e.g. Kenny and Joffres, 2008). Policy-makers have therefore searched for tools to make the decision-making processes on the allocation of health care resources more transparent and the final decisions less controversial, two goals hoped to be achieved by employing procedures such as HTA. Klein summarises this hope expressed by some of his colleagues in the following way:

[...] NICE represents an attempt to depoliticize decisions about who should get what [...]. Science and evidence, not political whims or media panics, would shape NICE's decisions and NHS practice: clinical and cost effectiveness would be the criteria when appraising technology and issuing guidance. Such at any rate was the theory when NICE was set up (Klein, 2010, pp. 389-393).

However, the ability of HTA processes to contribute to a 'de-politicisation' of the complex distributional choices in health care is increasingly questioned. For example, Landwehr argues, "[...] that the link commonly drawn between delegation to expert bodies and de-politicization of an issue [...] does not seem to hold [...]" (Landwehr, 2009, pp. 599-600). In the context of health policy this view suggests that commissioning independent and scientific expert bodies such as NICE and IQWiG with the task of recommending which pharmaceuticals should be made available on a national health care scheme does not solve the problem that recommendations might give rise to political opposition. As we shall see, the empirical results documented in this thesis support Landwehr's assertion in the sense that they show that the effect of the 'de-politicisation' of health care decisions appears to be questioned by policy-makers in England and Germany alike, leading to an increasing re-politicisation of the area.

In addition to the above, several authors question the ability of HTA processes to contribute to cost containment by identifying ineffective or wasteful health care interventions. By conducting HTAs health decision-makers can distinguish between cost effective and cost ineffective medical products, thus allowing for an exclusion of the latter in the health care benefit baskets of a given country. However, so-called disinvestments as a result from HTAs are still limited and observers such as Mossialos, Walley and Mrazek fear that HTA "[...] may become one of the drivers of health care costs" (Mossialos, Walley and Mrazek, 2004, p. 14) because more products are approved than denied.

Coulter and Ham (2000) point out that the debates surrounding HTA policies are not just marked by questions on whether it is an appropriate policy tool to address problems at hand, but also what format HTAs should take, how they should be carried out and conducted. Rather than presenting policy and decision-makers with an uncontested and rational tool for evaluating the effects of new pharmaceuticals, the methods used to that effect are varied and have given rise to disputes between experts. While HTA "[...] is a field of applied research that seeks to gather and synthesize the "best available evidence" on the costs, efficacy, and safety of health technology"

(Lehoux and Blume, 2000, p. 1083), the means to achieve this objective are varied and subject to contestation. Cost effectiveness analyses, cost-benefit analyses, measurement of outcomes in quality-adjusted life years (QALYs) and the efficiency frontier are examples of methods for the evaluation of pharmaceuticals<sup>7</sup>. They differ in terms of the focus they set and the input they require, thus making it likely that they differ also in terms of the outcomes they produce. Thus arises the methodological dimension of the controversies surrounding HTAs.

In addition to the political and methodological dimensions of HTA a third body of literature deals with the normative issues that priority setting and HTA give rise to (e.g. Norheim, 2002; Biller-Andorno, Lie and Meulen, 2002). Coulter and Ham argue that one of the lessons learnt from international rationing experience is the

[...] attention [drawn] to role of values in rationing. This is because the relative priority attached to different types of treatments or services [...] depends in part on the value attached to different outcomes (such as improving the quality of life as opposed to increasing the length of life). [...] The need to make these choices illustrates the ethical dilemmas involved in rationing and the moral basis of decision making (Coulter and Ham, 2000, p. 10).

Similarly, Kenny and Joffres (2008) assert that priority setting is essentially a normative and ethical process. Authors such as Biller-Andorno, Lie and Meulen have argued that coverage decisions based on HTA and evidence-based medicine threaten “[...] shared social values like equity and solidarity” (Biller-Andorno, Lie and Meulen, 2002, p. 261). The criticism emerges from their skepticism of the methods used for economic evaluations in HTA such as the use QALYs which they describe as a “[...] decision making procedure divorced from real life choices” (Biller-Andorno, Lie and Meulen, 2002, p. 269). Klein argues that the effect of rationing is that “[...] patients receive less than the best possible treatment” (Klein, 2010, p. 390), which makes the need for careful judgement and justification a pressing concern for decision-makers.

Holm (1998) and Landwehr (2009) assert that the above concerns arise from a second phase of rationing in which decision-makers have become increasingly aware that the introduction of HTA institutions does not preclude the necessity of making difficult value judgements. The body of literature on these concerns recognises that HTA gives rise to difficult value judgements, which in turn has led to a call for

---

<sup>7</sup> The differences between these methods are explained in chapter 5.

designing transparent and accountable decision-making processes in an effort to address them. Landwehr calls this development a case of “deciding how to decide” (Landwehr, 2009) while Daniels promotes a framework of “accountability of reasonableness” (Daniels, 2000), stipulating that it is easier to agree on fair processes than it is on content values that should guide decision-making.

The brief overview of the political/economic, methodological and societal/ethical dimensions of HTAs highlights the political and ethical salience of HTA processes. Empirically, they are illustrated by the medicines Kadcyła and Ticagrelor. The breast cancer drug Kadcyła, for example, was not recommended by NICE on the grounds that it was too expensive even though consultants and patient groups agreed that the clinical benefits from treatment with Kadcyła are substantial (BBC, 2014). Given the controversies and theories surrounding the effect of HTAs, the question arises what factors actually determine the outcome of pharmaceutical benefit assessments. In other words, do benefit assessment outcomes reflect political, economic, methodological, ethical and/or purely medical considerations? And what impacts on the relative balance between these considerations in different countries?

As the next section highlights, the extant literature on the factors that contribute to the outcome of pharmaceutical benefit assessments is limited in that it does not include in-depth comparative empirical studies of cases of benefit assessments in order to determine the factors that matter most within different national contexts. This thesis addresses this shortcoming. In doing so, it hopes that the contributing factors that are discovered during the course of the analysis give insights into the extent to which the political, economic, methodological, ethical and medical dimensions of HTAs play a role in different national contexts and how they shape the final outcome of HTAs. In this sense, the preceding overview of the controversies surrounding HTAs also serves as a preliminary framework for interpreting the results of the empirical analysis presented in chapters 6-8. It can only serve as a preliminary framework because the analysis is the result of a deductive process in which themes emerged during the analysis rather than a priori. Nevertheless, the controversies surrounding HTA provide an initial set of themes that may appear in the empirical data.

### 1.3. Gaps in Knowledge on HTA

Whilst the debates regarding the political, methodological and societal dimensions of HTA exhibit some detail, what is missing from the current literature is an empirical analysis of these dimensions and how they interact with each other. The opportunities for empirical research in this field are vast due the growing number of HTAs that are conducted and published on the websites of HTA bodies. The few studies on this question remain incomplete in the sense that they only explain variations in outcomes with reference to institutional differences and different statutory HTA criteria. The role of these factors notwithstanding, the extant literature on HTA lacks in-depth empirical examinations of whether these explanations offer an accurate account of what determines the outcome of pharmaceutical benefit assessments.

Kanavos, et al. (2010) examine the differences and similarities in appraisals and recommendations made by HTA agencies in six different countries (Canada, Australia, England, Scotland, Sweden and France). The results of their study are significant in that they show that HTA outcomes differed in more than half of the cases (Kanavos, et al., 2010, p. 2). The study also looks at a number of decision-making criteria that contributed to the final outcome and concludes that “[...] there are considerable disparities in the information required, interpretation of evidence, rigour of the appraisal process and stated motivations for listing or not listing drugs” (Kanavos, et al., 2010, p. 4). Similar conclusions can be found in Pomedli (2008) and Sorenson and Chalkidou (2012). Sorenson and Chalkidou (2012) include HTA objectives, processes, stakeholder involvement, assessment method and the application of evidence to decision-making as factors that impact on HTA outcomes and explain variations. However, while these studies make references to the political and ethical dimensions of HTA, their analysis remains focused on institutional features and methodological approaches rather than on analysing the interplay between the different factors in various cases. Moreover, they do not exhibit the same level of detail that is provided through the process-tracing approach contained in this thesis.

The above studies are important to the research question at hand because they highlight the empirical puzzle that motivates this research, namely the observation that the outcome of HTAs varies even if they are concerned with the assessment of the same medicine with largely the same evidence base. As outlined, the studies also provide some explanations for these variations. However, the explanations fall short of

demonstrating how different factors interact to produce certain outcomes. For example, while the studies acknowledge that institutional factors play a role in determining HTA outcomes, they do not demonstrate how these institutional factors might interact with ideational factors such as ethical considerations and value judgements. Moreover, while reasons for variation offer interesting insights, reasons for similarities are by and large not covered in the aforementioned studies. However, as highlighted in the empirical chapters of this thesis, the reasons for similarities also offer useful insights as there are a number of cases in which HTA bodies have reached the *same* conclusions despite considering *different* issues during the decision-making process. This is counterintuitive to what one might expect and offers an additional explanation for why a detailed empirical study on what determines the outcome of pharmaceutical benefit assessment is appropriate.

The next section outlines how this thesis is structured in order to address the research question of what determines pharmaceutical benefit assessments in health care systems that employ formalised HTA procedures.

#### **1.4. Outline of the Study**

The sections above placed the introduction of HTAs as a policy tool to aid health care decision-making in the context of a general reform trend that includes measures to ration and prioritise services in OECD countries. Even though HTAs are conducted on a number of the same pharmaceutical products with the same evidence base in different countries, existing research in the field has shown that the outcome of HTAs frequently varies from country to country. This empirical puzzle, along with the fact that current studies on HTAs are limited and fall short of providing a satisfactory account of this puzzle, gives rise to the following research question:

*What determines the outcome of pharmaceutical benefit assessments in health care systems that employ formalised HTA procedures?*

The next chapter presents the theoretical framework that underpins this research. The theoretical framework emerges from the literature on the role of policy paradigms and ideas in public policy. The focus on policy paradigms offers an attractive approach for addressing this thesis' research question because it allows for the consideration of a

multiplicity of factors that may play a role in determining the outcome of pharmaceutical benefit assessments. It also offers an opportunity to account for contextual, that is country-specific, factors that are likely to play a role since the characteristics of the pharmaceutical products that are analysed as case studies are controlled for and can therefore be ruled out as explanatory variables.

Paradigms are conceptualised as intellectual frameworks that determine which factors and issues are considered important in the process of a pharmaceutical benefit assessment. Based on the extant knowledge on how policy paradigms and ideas matter in the policy process, the underlying assumption is that HTAs, like other policies, can be understood as a reflection of specific health care and HTA paradigms. My theoretical argument is that, when HTAs are understood as a reflection of specific paradigms, these paradigms can help explain similar and dissimilar outcomes of benefit assessments.

In order to explain HTA outcomes with reference to paradigms, the paradigms have to be captured. Chapter 2 explains how I adapt Hall's (1989; 1993) and Majone's (1989) work in order to capture the dominant HTA paradigms in England and Germany<sup>8</sup>. It also elaborates on the importance of examining how paradigms operate in 'normal' practice. The concept of 'normal' practice derives from Kuhn's (1962) arguments on 'normal science' as the sphere in which paradigms operate. The underlying idea is that one can only hope to gain a better understanding of what a paradigm is and what its effects are if one looks at how it is established and articulated in the process of 'normal science' or, for the purpose of this thesis, the process of 'normal' decision-making practice in HTA.

The novel theoretical contribution lies in the fact that I apply theories of policy paradigms to an empirical puzzle other than policy change, which is the current preoccupation of authors in the field. Moreover, by extending Kuhn's (1962) concept of normal science to HTA I also emphasise the importance of examining a paradigm's features in 'normal' processes before looking at its role in policy change. In doing so, I demonstrate that the potential contribution of policy paradigms to explain empirical phenomena other than change in the policy process is underestimated, whereas their potential to explain policy change is overestimated. That is to say that the extant literature does not make it clear how paradigms can explain policy change if their

---

<sup>8</sup> See chapter 3 for an explanation of the case selection process.

distinctive features and how they are established in normal policy processes is not understood.

The third chapter outlines the research design of the thesis. The thesis represents a comparative case study with an embedded design. The reasoning behind this choice of study design as well as the selection of England and Germany as case studies and the choice for the embedded case studies of pharmaceutical products are explained in this chapter. Moreover, the methods for data collection and analysis are presented. The data analysis rests on the publicly available documents on pharmaceutical benefit assessments at NICE in England and the Federal Joint Committee (FJC) and IQWiG in Germany as well as on 23 interviews conducted with stakeholders who were involved in the decision-making processes in the ten cases that were analysed. The data was analysed using qualitative content analysis and process-tracing methods.

Chapter 4 offers a descriptive overview of health care, pharmaceutical policy and pharmaceutical benefit assessments in the two countries that were used as case studies in this research, namely England and Germany. It is important to understand the institutional and historical differences and similarities between the health care systems in England and Germany as they help explain some of the research findings that follow.

Chapter 5 presents an initial analysis of the policy paradigms of pharmaceutical benefit assessments in England and Germany by exploring the legislative and methodological frameworks that HTA systems are built on. In doing so, the chapter seeks to capture and outline the dominant HTA paradigms as they are laid out in statutory and methodological frameworks. It provides an introduction into what concepts, ideas and values guide pharmaceutical benefit assessments in the two countries. The discussion of the legislative and methodological basis of pharmaceutical benefit assessments helps with the interpretation of the results of the empirical chapters that follow. Essentially, at the end of the thesis, it allows for an assessment of whether the way benefit assessments are conceptualised in England and Germany impacts on how decision-making criteria are operationalised or whether we can observe divergences between theory and practice. Chapter 5 also contains a more detailed definition of the different methods that can be employed in carrying out pharmaceutical benefit assessments with reference to England and Germany.

The sixth chapter represents the first of three chapters that present and discuss the empirical findings. It outlines the six themes (see tables 3.5. and 6.1.) that arise from the empirical evidence, i.e. the consultation documents, stakeholder interviews and

others. Its focus is on the conceptualisation of evidence in England and Germany as the most important issue that determines the outcome of pharmaceutical benefit assessments. By analysing the decision-making processes in the cases of Cabazitaxel, Eribulin and Ipilimumab, all products for different types of cancer, the chapter offers a first insight into the finding that the way evidence is conceptualised and operationalised, and indeed what constitutes evidence in the first place, has a big effect on determining the outcome of pharmaceutical benefit assessments. It demonstrates that the issues that were considered in these cases reflect the values that are embedded in the paradigmatic constructs of pharmaceutical benefit assessments, outlined in chapter 5, but that divergences in values between countries do not necessarily lead to differences in benefit assessment outcomes. This is an important finding because it suggests that a similar outcome in the dependent variable (i.e. the outcome of the pharmaceutical benefit assessment) should not be equated with a similarity in judgements that were applied in a specific case.

In chapter 7 I discuss the operationalisation of the policy paradigms of pharmaceutical benefit assessments by tracing the decision-making processes in the cases of Fingolimod, Retigabine and Telaprevir. It is an extension of chapter 6 in that it also focuses on questions that are connected to the interpretation of evidence, which affect how policy paradigms of benefit assessments are operationalised. Using the examples of Fingolimod, Retigabine and Telaprevir I highlight that decision-makers make different judgement calls when it comes to questions such as the appropriateness of patient subpopulation or comparator products in benefit assessments. These judgement calls can lead to different outcomes of benefit assessments for the same product in different countries. The chapter concludes that the most important variable in determining the outcome of pharmaceutical benefit assessments is the ‘rules of evidence’ (Majone, 1989) that guide decision-making processes.

Chapter 8 outlines additional themes that arise from the empirical evidence. These themes are discussed as auxiliary variables because they did not arise in every embedded case study. Their relevance needs to be explored in future research in order to understand their meaning in a wider set of cases. The auxiliary variables discussed in chapter 8 include ‘public pressure’ in the form of media and public protests to recommendations by NICE in England. For Germany, the auxiliary variables include bargaining powers of different stakeholders and the conceptualisation of ‘patient relevance’. Whilst the external validity of the individual auxiliary variables cannot be

conclusively determined as they appeared to play a role in a limited number of cases only, the fact that they were different in England and Germany supports the finding that the policy paradigms under which pharmaceutical benefit assessments operate in the two countries differ from one another.

In the ninth and final chapter I draw together the research findings (summary in table 1.1.) and conclusions that arise from this study. There are eight research findings. The fact that five out of six themes that emerged from the data analysis were connected to questions around how scientific evidence should be interpreted gave rise to the first finding that ‘rules of evidence’ (Majone, 1989) play a substantial role in determining the outcome of pharmaceutical benefit assessments. Different rules of evidence help explain different outcomes. In the case of the epilepsy drug Retigabine, for example, different viewpoints by NICE and IQWiG on the appropriate comparator drug led to NICE recommending Retigabine (NICE, 2011a) whilst IQWiG did not conduct an assessment on the basis that the ‘wrong’ comparator was used by the pharmaceutical manufacturer (IQWiG, 2012).

The second finding presented in chapter 9 is that different rules of evidence are a reflection of different HTA paradigms, namely a paradigm of cost effectiveness in England and one of patient relevance in Germany. Within these paradigms certain concepts and criteria are emphasised more than others. Cost effectiveness is emphasised in England whilst patient relevance is emphasised in Germany. This gives rise to what may be labeled core and periphery concepts that decision-makers refer to when making decisions (research finding number 4). The core concepts in the English and German HTA paradigms are the thresholds, namely cost effectiveness and patient relevance, that need to be met in order for a pharmaceutical product to receive a positive assessment (research finding number 5). However, despite important differences between the English and the German HTA paradigm a significant finding is that these contrasting HTA paradigms do not necessarily lead to contrasting outcomes (research finding number 3). The case of Telaprevir that is discussed in chapter 7 demonstrates that evidence can be interpreted very differently and yet the outcome may be similar. This suggests that a similarity in the outcome of a pharmaceutical benefit assessment, i.e. in the dependent variable, should not be equated with a similarity in factors that were considered by the decision-making body. The process-tracing exercise conducted in the empirical chapters shows that the same assessment conclusion can be arrived at by the consideration of different issues. Finally, research findings 6, 7 and 8 (table 1.1.)

provide a better understanding of how paradigms are established in practice and of what kind of independent variables play a role in their operationalisation.

**TABLE 1.1. – Summary of Research Findings**

1. ‘Rules of evidence’ (Majone, 1989) matter
2. HTA paradigms take different forms
3. Contrasting HTA paradigms do not necessarily lead to contrasting outcomes
4. A core and a periphery of values exists
5. Thresholds are the expression of paradigms in normal decision-making
6. Paradigmatic coherence matters
7. Ideational and institutional variables matter
8. Paradigms can help explain empirical phenomena other than change

Chapter 9 also elaborates on the empirical and a theoretical relevance of the research findings. They are of empirical relevance because they contribute to a more comprehensive understanding of what determines the outcome of pharmaceutical benefit assessments and make the important point that a comparison of the outcome of the dependent variable alone does not fully explain differences and similarities. The thesis’ findings indicate that the definition and operationalisation of evidence questions, which are central to any HTA process, plays the most important role in determining the outcome of pharmaceutical benefit assessments. In the case of pharmaceutical benefit assessments, institutional variables such as the format that a given HTA body takes seem to be secondary to the ‘rules of evidence’ as an independent variable that affects assessment outcomes. In addition to the empirical contribution, the conclusion also summarises the significant theoretical contribution that this thesis makes by demonstrating that ideas in the form of policy paradigms matter, not just in times of policy change, but also in explaining the outcome of complex decision-making processes such as HTAs. Finally, I draw attention to some of the questions that this thesis was not able to address and in doing so I outline areas for future research.

## **1.5. Conclusion**

The introductory chapter provided the contextual background against which this thesis has to be understood. By reviewing the literature on HTA, it highlighted why HTAs in the form of pharmaceutical benefit assessments are a politically salient policy tool that gives rise to political, methodological and ethical controversies. By exploring the question of what determines the outcome of pharmaceutical benefit assessments in countries that employ formalised HTA procedures, this thesis seeks to make a contribution to the empirical knowledge on HTAs and on which factors play a role in determining their final outcome.

The overview of the literature on the dimensions of HTA illustrated that there are knowledge gaps of an empirical nature that are waiting to be addressed. These knowledge gaps relate to the lack of in-depth empirical studies that investigate how the different dimensions of HTA interact in certain cases which in turn demands an effort to connect theory and practice. Are the considerations that determine the final HTA outcome in pharmaceutical coverage decisions political, methodological or ethical in nature? Or is the nature of considerations dependent on the kind of product that is being assessed for coverage, i.e. does political or ethical salience take precedence over methodological and evidence-based considerations in some instances? These are some of the questions that the existing literature does not provide answers to. This thesis aims to fill this gap by examining the considerations that determine pharmaceutical benefit assessment outcomes in England and Germany.

In addition to the outlined empirical contribution, this thesis seeks to make a significant theoretical contribution by using policy paradigms as constructs to explain an empirical puzzle other than policy change. In doing so, I demonstrate that policy paradigms play a role in shaping the outcome of complex decision-making processes such as HTAs not just in times of change, but in times of relative stability and continued policy implementation. In the following chapter I explain the theoretical premise that my research is built on in more detail.

## **Chapter 2**

### **Theoretical Framework**

#### **2.0. Introduction**

This chapter outlines the theoretical approach that guides the empirical analysis in this thesis. The approach allows for an exploration of the factors that impact on the outcome of pharmaceutical benefit assessments. It is built around two main premises. Firstly, it emphasises the important role played by ‘evidence’ in health technology assessment (HTA) processes. This role will be operationalised with reference to what Majone (1989) labels the ‘rules of evidence’, that is the context-specific rules that determine, for example, which forms of evidence are prioritised over others. I make the theoretical argument that the ‘rules of evidence’, and other factors that emerge from the empirical analysis, are shaped by policy paradigms of pharmaceutical benefit assessments. This argument is connected to the second premise of the theoretical approach, namely that policy paradigms have an effect on the outcomes of pharmaceutical benefit assessment by determining what is or is not considered relevant in the assessment process. Drawing on Kuhn’s (1962) work on the connection between ‘normal science’ and paradigms, I emphasise that paradigms can be captured and understood by looking at how they are articulated in their routine application, e.g. during their application in pharmaceutical benefit assessment processes. Kuhn’s (1962) concept of ‘normal science’, Majone’s (1989) ‘rules of evidence’ and an adapted version of Hall’s (1993) concept of ‘policy paradigms’ represent the key theoretical principles that guide this research.

The following paragraphs introduce the theoretical reasoning that underpins the emphasis on the way in which rules of evidence are applied and the use of policy paradigms to understand the diverging interpretive approaches to HTA evidence in England and Germany. The concept of scientific evidence is the intellectual foundation upon which HTAs are built (e.g. Lehoux and Blume, 2000; Milewa and Barry, 2005). An assessment of a pharmaceutical product or medical intervention stands and falls by the availability and quality of evidence to prove effectiveness, cost effectiveness, additional benefit and other given outcomes one hopes to measure in the process. Given

that ‘evidence’ characterises the core of HTAs<sup>9</sup>, regardless of the institutional or national setting, it is reasonable to hypothesise that questions of evidence are a major contributing factor to the outcomes of pharmaceutical benefit assessments. This thesis thus begins with the theoretical assumption that ‘evidence’ is a key variable in determining the outcome of pharmaceutical benefit assessments in countries that employ formalised health technology assessments.

The above assumption is a solid theoretical starting point, but it does not explain why *prima facie* the same evidence on a pharmaceutical product does not, as a matter of course, lead to a similar assessment of the given product in different countries. Put differently, it does not explain why the same evidence leads to similar outcomes in some assessments and dissimilar ones in others. The availability and quality of evidence undoubtedly plays a role because it is the intellectual core of HTAs, but it cannot be the sole determining factor in explaining the outcomes of pharmaceutical benefit assessments. The fact that assessment outcomes are dissimilar in some cases but similar in others despite the availability of the same evidence base, suggests that the ‘evidence variable’ alone does not adequately explain the empirical puzzle and, that there is an important piece missing in solving the puzzle.

The empirical puzzle indicates there must be factors other than the mere presence or absence of good-quality evidence that explain why benefit assessment outcomes vary between countries. The evidence base underlying the ten case studies of pharmaceutical benefit assessments in this thesis is controlled for, i.e. the same randomised controlled trials (RCTs) and other clinical data are available for all ten cases due to the global nature of clinical trials. However, the empirical analysis of the decision-making processes in these ten cases indicates that HTA organisations in England and Germany attach different meaning to different types of evidence and prioritise clinical and other outcomes differently. This suggests that different *rules* are applied to the same evidence, which can result in contrasting interpretations of the evidence. It seems that it is the *interpretation* of the evidence, rather than any inherent characteristic of the evidence itself, that plays a key role in determining the outcome of pharmaceutical benefit assessments. According to Van Der Wilt and Reuzel: “[...] it is important to be aware of the fact that any HTA is likely to be constrained by normative

---

<sup>9</sup> See definitions and discussions on HTA in chapters 1 and 5.

considerations, determining those facts to which we will turn our attention” (Van Der Wilt and Reuzel, 1998, p. 352).

The previous remarks explain the emphasis on the *interpretation* of evidence as a key to approaching this thesis’ research question. Evidence questions are framed according to different criteria and values. This impacts on what counts as evidence in the first place and whether the evidence is deemed satisfactory. Biller-Andorno, Lie and Meulen argue that “[...] moral values often play a hidden role, not only in the production of ‘evidence’, but also in the way this evidence is used in policy making” (Biller-Andorno, Lie and Meulen, 2002, p. 261). The empirical puzzle therefore demands a theoretical approach by which one can discern and explain the rules of evidence that impact on a comparable, i.e. identical, evidence base in the form of RCTs and other global trial data, in a way that explains both converging and diverging assessment outcomes.

The thesis employs the concept of policy paradigms to explain both converging and diverging pharmaceutical benefit assessment outcomes between countries. Paradigms, henceforth conceptualised as intellectual frameworks that determine which issues are considered important and relevant in the assessment process, help explain why a number of features of an evidence base attain different meanings in different HTA systems and how this may lead to divergent HTA outcomes. Paradigms give rise to different rules of evidence and in doing so, paradigms function as guides for which pieces or aspects of evidential information will be considered relevant by decision-makers. Paradigms may also influence stakeholders’ positions and views in assessment processes, which may in turn impact on the final outcome of a given assessment. Understanding what determines the outcomes of pharmaceutical benefit assessment with reference to a policy paradigm framework is a two-step process in which, as a first step, both the paradigm and the rules of evidence have to be identified and, as a second step, the operationalisation thereof has to be examined.

What follows in the preceding sections is an overview of how policy paradigms have been conceptualised and employed as a theoretical construct in the literature on policy paradigms, ideas and the history of science. This includes an outline of the way in which policy paradigms are operationalised in this thesis as well as an outline of the main variables that emerge from the literature on policy paradigms. The chapter then continues by providing a brief overview of some of the ways in which challenges around questions of evidence are developed and explained in different branches of

literature such as the literature on evidence-based policy and policy analysis. This serves the goal of illustrating that evidence acquires different meanings in different contexts, a phenomenon that Majone (1989) associates with different ‘rules of evidence’ in policy systems. Finally, a brief excursion is taken to look at some of the alternative theoretical models that were considered in the research process and to explain why these models were not pursued.

## **2.1. Policy Paradigms**

The roots of the use of paradigms to explain a myriad of phenomena are frequently accredited to Kuhn (1962), who introduced ‘paradigms’ as overarching frameworks of ideas and worldviews to explain change and stability in scientific research within communities of scientists over the centuries. In exploring the history of science and scientific revolutions, Kuhn (1962) concluded that within a community of scientists paradigms influence which problems are defined as problems worth solving and, which methods and theoretical assumptions are deemed appropriate in trying to solve them. Kuhn originally defined ‘paradigms’ as “[...] universally recognized scientific achievements that for a time provide model problems and solutions to a community of practitioners” (Kuhn, 1962, p. viii). They give rise to “[...] some accepted examples of actual scientific practice – examples which include law, theory, application and instrumentation [...]” (Kuhn, 1962, p. 10). In other words, paradigms are more than overarching intellectual frameworks of ideas and worldviews, they also give rise to accepted rules, methods, theories and instruments to approach scientific puzzles. Scientific revolutions and change occur when the dominant rules, methods and theories are no longer adequate in explaining the problems that a scientific community is trying to solve, for example when anomalies in one’s theories arise from new discoveries that need to be explained (Kuhn, 1962, pp. 52-53).

The concept of ‘paradigms’, or ‘policy paradigms’, is the subject of an expanding strand of literature in political science and public policy, which emphasises the importance of ideas, values and beliefs in explaining the policy process and policy change. Theories on policy paradigms emerged from a dissatisfaction with another theoretical stream, namely that of new institutionalism, to explain institutional development and especially change (Béland and Cox, 2011a). This dissatisfaction gave rise to what Béland and Cox (2011a) refer to as “ideational approaches” – which

include policy paradigms - to the policy process. There is now a growing body of literature that suggests that ideas play a role in shaping the policy process (e.g. Hall, 1993; Campbell, 2002; Poteete, 2003; Menaheim, 2008; Béland and Cox, 2011).

A consistent definition of 'ideas' is lacking in the ideational literature. According to Poteete: "Ideas, broadly understood, encompass everything from normative and ontological beliefs to perceptions about the disposition of other actors to understandings of causal relationships" (Poteete, 2003, p. 531). Béland and Cox define ideas as "causal beliefs [...], products of cognition" (Béland and Cox, 2011a, p. 3), which provide guides for action. While the lack of a consistent conceptualisation of 'ideas' is acknowledged as a shortcoming, the issues that arise from it are secondary for the purpose of this thesis and are not expanded upon any further. For the purpose of this thesis, ideas are conceptualised as distinctive concepts related to one another in a logical framework of decision and analysis. These distinctive concepts can be normative or empirical assumptions about policy problems, goals and tools and they are embedded in HTA policy paradigms at any given time. As we shall see, an example of the role played by ideas is the distinctive and differing conceptualisation of evidence in HTA processes in England and Germany. While NICE recognises patient views as evidence in their own right, IQWiG does not consider views brought forward by stakeholders as evidence. This demonstrates a distinctive idea, a distinctive understanding, about the concept of evidence. For the theoretical argument of the thesis, an understanding of the concept of policy paradigms and the operationalisation thereof is more important than a precise definition of 'ideas'. Indeed, limiting the theoretical approach to a precisely defined outline of what an 'idea' is, would have run the risk of missing important ideational factors in the empirical analysis because a too narrowly specified definition of 'ideas' might have prevented it.

Despite some differences in the definition and operationalisation of paradigms in the public policy literature, common and recurring themes can be identified. These include an understanding of paradigms that denotes a common view of the world, be it the world of a scientific community or the world of a wider policy community (e.g. Kuhn, 1962; Hall, 1993; Poteete, 2003; Kay, 2007; Béland and Cox, 2013). They also include an assumption that values and beliefs are strongly embedded in paradigms and that, taken together, the worldview, values and beliefs give rise to a limited set of policy problems, options and tools in any given area (e.g. Hall, 1993; Béland and Cox, 2013). Paradigms are conceptualised as intellectual constructs that determine what is or is not

considered possible, viable, acceptable or desirable in a given field. As outlined in more detail in a later section of this chapter, Kuhn's (1962) work on the history of science and the role of paradigms therein had an influence on the conceptualisation of policy paradigms in this manner. The central role of paradigms as a limiting, boundary-defining, intellectual framework plays a major part in this thesis as it helps explain what is considered important in a given pharmaceutical benefit assessment process, i.e. what gets defined as 'evidence' in the first place, why notable differences in the rules of evidence in HTAs exist and to what extent stakeholder positions are permitted and considered during the process of assessment.

Arguably the most influential research on the role of policy paradigms is by Hall (1989; 1993). In line with Kuhn's (1962) analysis of paradigms, Hall employs the concept to explain how policy change occurs when issues arise that can no longer be explained or solved by using familiar policy options and tools. Underlying Hall's thinking is a conceptualisation of the policy process as:

[...] a process that usually involves three central variables: the overarching goals that guide policy in a particular field, the techniques of policy used to attain these goals, and the precise settings of these instruments (Hall, 1993, p. 278).

Hall argues that the three central variables outlined above gain meaning and content in the context of policy paradigms:

[...] policymakers customarily work within a framework of ideas and standards that specifies not only the goals of policy and the kind of instruments that can be used to attain them, but the very nature of the problems they are meant to be addressing [...] this interpretive framework is a policy paradigm (Hall, 1993, p. 279).

In other words, a policy paradigm acts as an umbrella of ideas, values and policy options under which policy actors engage and develop policies. Policy change occurs when there are alterations in policy paradigms, brought about by external or internal factors, that open up new policy options.

The function of policy paradigms as an umbrella of ideas, values, policy goals and options has been taken up by other authors in the field (e.g. Menaheim, 1998; Bergeron and Kopp, 2002; Kay, 2007). For example, in an article on changes in public health insurance policies in Australia, Kay argues that:

Policy paradigms shape and structure how problems are framed; the imperatives for reform and the reform proposals offered; they are a relatively coherent set of assumptions about political, social and economic institutions and a policy road map of how to reform existing or how to introduce new programs. They contain a set of cognitive and normative maps or frames that orientates actors within a policy sector, defines the problems, the goals and the means of achieving these (Kay, 2007, p. 583).

Similar views on policy paradigms can be found in case studies on a range of policy areas, the number of which is growing. Policy paradigms have been used as theories to explain phenomena in the areas of pension policies (e.g. Melo, 2004; Orenstein, 2013), water policy (Menaheim, 1998), property rights policies (Poteete, 2003), drug abuse policies (Bergeron and Kopp, 2002), public health insurance policies (Kay, 2007) and economic policy (Hall, 1989). Most of these studies are case studies that are confined to a particular country, e.g. water policy in Israel (Menaheim, 1998) or property rights policies in Botswana (Poteete, 2003). Common to these case studies is a conceptualisation of policy paradigms as a framework of ideas that influence policy problems and solutions. Frequently, they also examine the role of institutions as vehicles or obstacles to the implementation of ideas (e.g. Hall, 1989; Poteete, 2003). For example, Hall (1989) attributes the move towards Keynesian policies in post-war Great Britain to institutional structures that lent themselves well for Keynesian ideas. He asserts that: “[...] ideas acquire force when they find organisational means of expression” (Hall, 1989, p. 275).

Hall’s (1989) preceding assertion is taken up in this thesis in the sense that HTA institutions such as the National Institute for Health and Care Excellence (NICE) and the Institute for Quality and Efficiency in Health Care (IQWiG) are understood to represent the respective HTA paradigms in England and Germany. The aim of this research is to identify the HTA paradigms in England and Germany and to compare and contrast how they impact on assessment outcomes. In light of this aim, I assume that national HTA paradigms become established in HTA bodies and I underline this point by examining how the main ideas of a given paradigm are applied in practice during the decision-making processes of HTA. I examine how these paradigms are being implemented *post-policy change* – the change from a situation in which there were no HTAs to a situation in which HTAs have become an established feature in health care decision-making – and how this affects the final outcomes of benefit assessments. Institutional variables are considered in the analysis, but they are not at the forefront

because they give little insight into how rules of evidence are operationalised within different HTA policy paradigms. Put differently, identifying that an idea has been institutionalised does not give insight into how this idea is operationalised in a ‘normal’ decision-making process. The reference to the concept of a ‘normal’ decision-making process is based on Kuhn’s (1962) concept of ‘normal science’, which is explained in more detail in the next section.

Finally, the body of literature on policy paradigms makes an additional important theoretical contribution. It highlights that a number of different paradigms may exist in a policy area at any given time. For example, Bergeron and Kopp (2002) juxtapose the “harm reduction paradigm” against the “psychoanalysis paradigm” to explain the perceived uniqueness of French drug abuse policies and the changes that have been introduced more recently. Similarly, Menaheim makes a distinction between two water policy paradigms in Israel between 1948-1997, namely a paradigm “[...] of expanding water resources and agricultural production [...]” and “[...] a paradigm of priority of agricultural expansion over water conservation” (Menaheim, 1998, p. 283).

The assumption that a number of different paradigms exist in a policy area suggests that any research that is built on the centrality of policy paradigms as a theoretical model requires an awareness of the currently existing policy paradigms and their main principles. Moreover, it requires an understanding of the currently *prevailing* paradigm in a policy field. For the purpose of this thesis, the prevailing HTA paradigm will be labeled the ‘dominant’ or the ‘emerging’ paradigm to denote its importance. In approaching the research question the task of identifying the emerging dominant HTA paradigm is comparatively more important than identifying other, perhaps auxiliary, paradigms. This is because I argue that the outcomes of pharmaceutical benefit assessments are affected by divergent HTA policy paradigms in different countries; underlying this argument is the assumption that *dominant* HTA paradigms emerge in the process of HTA decision-making and that these impact on benefit assessment outcomes. Whilst the empirical analysis does provide some insight on auxiliary paradigms, the focus of the thesis is on exploring a) how the *dominant* HTA policy paradigm can be described, b) how it is operationalised during the ‘normal’ pharmaceutical benefit assessment process and c) how this operationalisation affects the final outcome of assessments.

In summary, the literature on paradigms and policy paradigms gives rise to three key findings. Firstly, it ascribes an important role to ‘ideas’ in the policy process (e.g.

Béland and Cox, 2011; Kurzer, 2013). Ideas are embedded in paradigms that function as an umbrella construct under which certain policy options, or outcomes, are possible while others are not. Change can occur under a given umbrella, but the impact of change depends on the kind of policy challenges that it needs to respond to (Hall 1993; 1989). The more significant the challenges, the more likely it is that an old paradigm gives way to a new paradigm (Hall, 1993). Secondly, institutions matter in the policy process. They matter because they are a reflection of certain policy paradigms that may or may not have changed over time (Radaelli, 1995). And thirdly, a number of policy paradigms can exist in a given policy area at any given time (e.g. Menaheim, 1998; Bergeron and Kopp, 2002) and it is for the researcher to identify these paradigms in order to distinguish between emerging dominant and auxiliary paradigms.

## **2.2. Shortcomings in the Extant Literature**

The majority of researchers who employ ideational approaches generally, and policy paradigm approaches specifically, are concerned with explaining policy change. In doing so, and in line with Kuhn's (1962) original thoughts on the subject, the focus is on paradigm shifts. Policy change is perceived as the dependent variable whilst policy paradigms, and the changes therein, are independent variables. Policy change is described as a shift in worldviews and perceptions that opens up previously impossible policy routes and possibly "[...] usher[s] in a new era [...]" (Hall, 2013, p. 191). The preoccupation lies with shedding light on the reasons why one paradigm became dominant over another and the circumstances that make transitions from one dominant paradigm to another possible (e.g. Bergeron and Kopp, 2002; Kay, 2007; Menaheim, 2008). Here, Hall (1993) distinguishes between first order (affecting the instrument settings), second order (affecting the instruments) and third order (affecting the overarching goals) change in the policy process. Hall (1993) argues that the likelihood of true paradigmatic change, i.e. pursuing a policy path that was previously closed off, depends on which of the three policy process variables is targeted by a novel idea, how this idea sits with the current policy paradigm and the institutional characteristics of a system.

The preoccupation with policy change represents a shortcoming in the current literature that employs policy paradigms as theoretical models to explain a variety of empirical phenomena. It is a shortcoming because it has led to an almost exclusive

focus on empirical phenomena of change in policy processes. In turn, this has led to a comparative neglect of the equally interesting question of how emerging dominant policy paradigms shape day-to-day outcomes in given policy areas such as HTA policy. In other words, what role do policy paradigms play *post*-policy change, in the implementation phase of the policy cycle? How are they operationalised and put to use?

The aforementioned questions are comparatively under-researched in the literature. The previous discussion of case studies that employ policy paradigms as theoretical constructs has shown that there is a tendency to emphasise the role of paradigms *prior* to policy implementation. In other words policy paradigms are used to explain the outcomes of the policy-making process. However, the theoretical premise of this thesis is that policy paradigms can also contribute to a better understanding of how policy is implemented, and to what effect, by looking at how paradigms operate when policy is put into practice. Theories on policy implementation have emphasised the need to examine how policy gets translated into practice, and perhaps transformed through the involvement of actors such as professionals and bureaucrats (e.g. Dunleavy, 1981; May and Winter, 2009; Honig, 2006), which in turn affects how policy is adapted at a later date. Policy paradigms can offer an additional angle in the analysis of implementation.

However, it is not just against theories on policy implementation that the use of paradigms is worth exploring. It is also worth exploring because the traditional view of the policy process as consisting of several distinct phases of policy-making has been criticised. Several authors have argued that this view of policy-making is oversimplified and that, in reality, where policy-making ends and implementation begins (and vice versa) is more difficult to discern, with one affecting the other (e.g. Sabatier and Jenkins-Smith, eds., 1993). In line with this view, this thesis offers an additional theoretical perspective by examining the role of paradigms in policy implementation as an extension of the fluid process of the policy process.

The problem that arises from the current focus in the literature on policy paradigms is that it implies that they are operationally static or insignificant during times in which policy changes, i.e. significant paradigm shifts, are not occurring. As a result there is little empirical understanding of the effect and characteristics of paradigms on 'normal' decision-making processes. For the purpose of this thesis decision-making processes are conceptualised as the argumentative processes that lead to decisions on the benefits of pharmaceutical products in the context of HTA. They are

the processes that occur when HTA policy in a given country is implemented and refined. The emphasis on *normal* processes is derived from Kuhn's insights on the importance of 'normal science' for capturing and understanding paradigms (Kuhn, 1962, pp. 10-11). Kuhn introduces the term 'normal science' as a concept that is closely associated with the dominant paradigm in a given scientific community. Kuhn's delineations on 'normal science' are frequently neglected in the public policy literature on paradigms, but the following paragraphs show that the conceptualisation of 'normal science' as processes in which paradigms are articulated and refined is useful in capturing the essence of paradigms in public policy.

In a postscript to a later edition to his original piece of work, Kuhn (1969) distinguishes between a paradigm as consisting of beliefs and values shared by a certain community and a paradigm as encompassing more practical aspects of scientific research, for example appropriate methods and models in solving a problem (Kuhn, 1969, p. 175). Over the years, Kuhn's delineation of the term 'paradigm' has received much attention, some of it critical when it comes to the appropriateness of defining and using the term in the way he suggests. A discussion of these criticisms goes beyond the purpose of this section, let alone the aim of this thesis as a whole. However, Kuhn's (1962) contributions with regards to what he labels 'normal science' are helpful in underlining this thesis' theoretical approach and contribution.

According to Kuhn:

[...] 'normal science' means research firmly based upon one or more past scientific achievements, achievements that some particular community acknowledges for a time as supplying the foundation for its further practice (Kuhn, 1962, p. 10).

When these achievements provide a sufficient number of scientific problems to solve and when they are more convincing than their competitors in offering routes to explanations of these problems, then they can be referred to as paradigms (Kuhn, 1962). In other words, in Kuhn's view the terms 'paradigm' and 'normal science' lie closely together. Normal science can be thought of as *paradigms in operation* (Kuhn, 1962, p. 11), that is an expression of the way paradigms are developed, established and put into practice.

Kuhn also refers to 'normal science' as paradigm-based research (Kuhn, 1962, p. 25), that is research traditions that consist of laws, theories, applications and

instrumentations (i.e. methods) that guide the work of scientific communities. Kuhn illustrates that these traditions of ‘normal science’ or paradigm-based research do not come to exist overnight, but that they are result of a periods of development and adjustment within branches of science by which paradigms of how to conduct research become ‘normalised’. Over time, the community of scientists comes to agree on what facts are relevant to their discipline, how these facts relate to dominant theories and how research should be conducted to further articulate those theories (Kuhn, 1962, pp. 32-33). In the process of articulating research paradigms, rules, for example on the appropriateness of certain methods over others, emerge. Here, Kuhn asserts that the study of normal science and the rules contained therein provides insights into the given paradigm one is studying: “The study of normal-scientific traditions discloses many additional rules, and these provide much information about the commitments that scientists derive from their paradigms” (Kuhn, 1962, p. 40).

Kuhn asserts that different rules between research paradigms can help explain why similar problems and solutions are viewed differently by different scientific communities. He illustrates his point with the following example:

An investigator who hoped to learn something about what scientists took the atomic theory to be asked a distinguished physicist and an eminent chemist whether a single atom of helium was or was not a molecule. Both answered without hesitation, but their answers were not the same. For the chemist the atom of helium was a molecule because it behaved like one with respect to the kinetic theory of gases. For the physicist, [...], the helium atom was not a molecule because it displayed no molecular spectrum. Presumably both men were talking of the same particle, but they were viewing it through their own research training and practice (Kuhn, 1962, pp. 50-51).

This example highlights the importance of normal science, paradigms and rules to the framing of problems and solutions. It is key to understanding this thesis’ theoretical approach. Paradigms can help explain why answers to similar questions, namely the question of what constitutes the benefits of a pharmaceutical product, are frequently different in different countries. Assessment outcomes vary as a result of paradigm differences and, as Kuhn points out, such paradigm differences can be “consequential” (Kuhn, 1962, p. 51). In health policy they are consequential in the sense that they lead to differences in the access to new medicines that is covered by public health care systems.

By emphasising the connection between normal science and paradigms, Kuhn (1962) makes the point that paradigms should not be thought of as static constructs, but

rather as constructs that get developed during the processes of normal science in which certain questions are considered worth pursuing by a given scientific community. He argues: “In a science, [...], a paradigm is rarely an object for replication. Instead, like an accepted judicial decision in the common law, it is an object for further articulation and specification [...]” (Kuhn, 1962, p. 23). He further argues that paradigms promise to contribute to the solution of problems and that normal science is the “actualization of that promise” (Kuhn, 1962, p. 24). It is this latter thought, the ‘actualization of the promise’ of a paradigm, that influenced the theoretical focus and methodological approach of this thesis. Essentially, Kuhn argues that paradigms get articulated when put into operation and that during this process of operation they are further refined, i.e. their promise is actualised (or not, depending on the strength and security of the paradigm). This suggests that one cannot hope to capture, let alone understand, the elements of a paradigm without examining how it is operationalised and refined in ‘normal’ practice. Hence the focus of this thesis is on analysing how policy paradigms of HTA are operationalised in England and in Germany and how this affects the final outcome of the assessments.

The fact that Kuhn’s argument stems from an analysis of what he refers to as scientific revolutions does not make it less applicable to the study of public policy generally and health policy specifically. This is because one of his main points, namely that paradigms are refined in the process of engaging in normal science, is transferable to many other subject matters. One can learn about the impact of paradigms on empirical phenomena such as the outcome of HTAs not just by assuming that they are influential, but by examining how they are established in ‘normal’ practice and how this ‘normal’ practice leads to manifestations or alterations in a given paradigm. Building on Kuhn’s concept of normal science as paradigm-based research I argue that HTA decision-making is essentially paradigm-based and that the identification of HTA paradigms can help explain variations and similarities in HTA outcomes.

As we shall see in chapters 6-8, the analysis of the ten case studies of pharmaceutical benefit assessments demonstrates that the influence of HTA paradigms is nuanced. Different paradigms do not, as a matter of course, lead to different HTA outcomes, but they help explain the different conceptual lenses through which, for example, pieces of evidence are viewed. By looking at how the paradigms are applied in normal practice, one can learn about which elements undergo further definition, which ones are open to flexible interpretations and which ones represent minimum ‘thresholds’

that have to be met in order to fit with the paradigm. Put differently, HTA paradigms undergo further definition and specification beyond their formulation in statutory or methodological frameworks and this occurs during their routine application in specific decision-making or policy-making processes.

In summary, the main shortcomings of the extant literature on paradigms in public policy is the lack of studies on how paradigms are operationalised in periods other than policy change. Currently, policy paradigms are used as concepts to explain policy change without an in-depth elaboration on the characteristics of the paradigms and, more importantly, how they have become articulated and specified in normal practice. The implications of this shortcoming should not be underestimated because the question arises how one can seek to understand the role of paradigms in policy change without fully understanding the way they are (were) put into practice in the first place. Understanding the operation of policy paradigms is undoubtedly more arduous than describing their role in policy change. However, ultimately it might lead to a deeper understanding of paradigms and perhaps a more precise definition of what they are in the context of public policy-making.

The merits of the above approach have been highlighted in the previous paragraphs. However, a note of caution is necessary in that ‘normal’ decision-making in HTA may be more challenging to identify than the ‘normal science’ of scientific communities that Kuhn (1962) refers to. This is in part because HTA processes include a number of different actors, with different interests and from different disciplines. However, despite this complexity the empirical analysis contained in this thesis shows that patterns, rules and methods do exist in the ‘normal’ HTA decision-making processes and that these can be viewed as a reflection of different paradigms. Whether or not these patterns, rules and methods coincide with the belief systems (Sabatier, 1988) of one community of stakeholders rather than another is secondary for the purpose of this thesis’ research question. As a first step, this thesis seeks to understand what determines the outcomes of pharmaceutical benefit assessments in different countries and how paradigms can be used to explain these outcomes. The second step, i.e. a closer examination of whether the paradigms align with those of certain stakeholder communities or whether they are indeed ‘shared paradigms’, is a step that goes beyond the scope of this research and would have to be undertaken as a separate research endeavour in the future.

In conclusion, this thesis does not seek to explain why policy change occurred. Instead, it seeks to explore how a policy paradigm, i.e. the *dominant* HTA paradigm in a given country, becomes established in a normal every-day decision-making process and how this affects the outcome of the process. In line with the extant literature in the field, the function of policy paradigms as an umbrella construct that encompasses a set of specific policy options and tools is at the heart of this research, but rather than examining what happens to policy paradigms in periods of transition, or how they explain these transitions, I examine how they matter in periods of perceived stability and how they impact on a politically salient case of decision-making in health care policy, i.e. that of pharmaceutical benefit assessments.

In comparison to other policy paradigm approaches, the principles contained within policy paradigms still feature as the independent variables for the purpose of this study. However, in contrast to other case studies, the dependent variable is not policy change, but rather the outcome of decisions on the benefits of pharmaceutical products. My theoretical contribution lies in the expansion of the use of policy paradigms as theoretical concepts to explain empirical puzzles other than change. The theoretical insights from this research are strengthened by the comparative cross-national case study approach that is employed as this approach discloses different paradigms and how they operate under ‘normal’ circumstances. By identifying the HTA paradigms in England and Germany and comparing and contrasting how they are articulated in normal benefit assessment processes on the same pharmaceutical products, a deeper understanding of the function of paradigms as fluid constructs that influence outcomes is gained.

The next section outlines how the above findings are adapted and developed in an effort to explain what determines the outcome of pharmaceutical benefit assessments in countries that employ formalised HTA procedures.

### **2.3. The Policy Paradigms of Pharmaceutical Benefit Assessments: A Conceptual Framework**

From the discussion of the literature on policy paradigms, one can conclude that different paradigms, or ways of approaching a problem, exist in policy fields. This suggests that different paradigms exist within national and international HTA decision-making and it might help explain what determines the outcome of pharmaceutical benefit assessments. In order to explore the validity of the assumption that different policy paradigms exist within the HTA policy arena, it is useful to recall how one of the professional networks dedicated to the promotion of HTA defines the concept. Here, the general idea of HTA is defined as:

[...] a multidisciplinary process that summarises information about the medical, social, economic and ethical issues related to the use of a health technology in a systematic, transparent, unbiased, robust manner. Its aim is to inform the formulation of safe, effective, health policies that are patient focused and seek to achieve best value (EUnetHTA, 2014).

The above definition of HTA implies that decision-makers have a range of issues and tools to choose from when it comes to designing HTA policies. How the general ideas of HTA are articulated and developed in a given national context is crucial with regards to understanding HTA paradigms. Decisions have to be made with regards to what counts as medical, social, economic and ethical issues in a given country and whether there are boundaries to these issues. Decisions also have to be made with regards to which methods are considered most appropriate to provide a transparent and unbiased assessment. Finally, an agreement has to be found on what constitutes a ‘focus on the patient’ and ‘best value’, especially because it is easy to conceive of a situation in which patient focus might not sit well with the idea of value for money and vice versa. For example, ‘patient focused’ policies may give rise to the view that medicines that extend a patient’s life by 3-4 months at the end of his/her life should be available and covered by the public health system, whereas a ‘best value’-approach would suggest that 3-4 months of additional lifetime gained does not justify the high costs of these medicines and that the money might be better spent elsewhere.

The above scenario exemplifies a common dilemma in health care priority setting. That is the dilemma of making judgements on issues such as the ‘right’ care pathways and ideology on access to health care to apply when these might be in conflict

with each other (e.g. Lehoux and Blume, 2000; Van Der Wilt and Reuzel, 1998). Resolving this dilemma entails making difficult choices and the outcome of these choices is, for the purpose of this thesis, conceived of as a reflection of different policy paradigms. Just as Kay argues that “[...] universal public health insurance has a different meaning in a paradigm with patient choice at its core compared to one based on equality of access to service” (Kay, 2007, p. 583), I argue that HTA generally, and HTA evidence specifically, has a different meaning depending on the core features of a paradigm. In the context of the previous example, I would argue that the presented dilemma can be understood as a contest between a ‘patient focused’ paradigm and a ‘best value for money’ paradigm.

In the policy field of HTAs one is faced with a situation in which, from a single definition of HTA, potentially multiple sub-paradigms emerge. The relative priority that is given to issues contained in sub-paradigms depends on ideational and institutional factors that are specific to national contexts. As Littlejohns, et al. point out: “Even explicit and seemingly “scientific” criteria such as clinical and cost effectiveness are embedded in views about, for example, the value of different health states” (Littlejohns, et al., 2012, p. 286). Given the complexities of HTAs, it is reasonable to assume that a variation in HTA paradigms between different countries exists, a variation that depends on the issues, criteria and methods – the rules of evidence - that are put at the centre of the assessments. These issues, criteria and methods are in turn shaped by the values and ideas that are embedded in the health care context, i.e. the overarching health care paradigm. Applying a policy paradigm framework to the HTA policy area thus entails the identification and analysis of paradigms at two levels that differ in terms of the range of fields for which they are relevant.

At a first level one can identify paradigms that are limited to the specific policy area that is the subject of one’s study, e.g. drug abuse policy, inflation policy or, as in the case of this thesis, HTA policy. Within these policy areas one can describe currently dominant paradigms that give rise to problems, solutions and policy tools. These are what Hall labels “[...] narrower paradigms that often dominate specific fields of policy” (Hall, 2013, p. 191). The problems that are tackled and the tools that are used are a reflection of what Kuhn (1962) labels ‘normal science’ and they provide insights into how paradigms operate in normal routine practice. For example, in relation to evidence questions in HTA they help explain why HTA bodies may have divergent views on the same evidence on the same pharmaceutical products.

The specific ‘narrow’ paradigms are usually not applicable to other policy areas such as elderly care policies or drug abuse policies for example. However, they are shaped by broader paradigms of values and worldviews that are operational on a second analytical level. HTA, elderly care and drug abuse policies operate within wider health, social care, and public health paradigms that need to be considered if one seeks to understand the narrow paradigms of specific fields of policy (Hall, 2013). For example, Menaheim (1998) illustrates that water policy paradigms are closely connected with agricultural policy in Israel and explains it with reference to the overarching Zionist ideology – or paradigm - which ascribes a crucial role in Israel’s nation-building efforts to agricultural policy measures. When applied to health care, the above remarks suggest that the dominant HTA paradigm will also be a reflection of the overarching health care paradigm. That is to say that core features such as utilitarian or individualistic approaches to health care are likely to give rise to different HTA paradigms.

For the purpose of this thesis, the narrow HTA paradigms emerge from statutory frameworks and methodological guidelines that HTA bodies are required to adhere to. However, as highlighted previously, their precise elements and articulation can only be understood by analysing how they are applied in practice during the normal pharmaceutical decision-making processes, which explains why the empirical focus of this thesis is on the ten case studies of assessments. The broader health care paradigms emerge from the statutory and constitutional pillars of health care. However, while the broader health care policy paradigms will be identified, they are not the focus of this thesis. It is important to be aware of them and to identify them in order to understand to what extent the narrow HTA paradigms are a reflection of broader paradigms, but for the purpose of answering the research question, the focus is on analysing how different *HTA* paradigms are operationalised. The aim is to examine how different rules of evidence, stakeholder views and institutional characteristics that emerge from HTA paradigms impact on the outcomes of pharmaceutical benefit assessments.

In the next section I sketch how I adapt the independent variables that arise from the literature on policy paradigms before offering an overview of the crucial role of evidence and rules of evidence (Majone, 1989) in the way policy paradigms are operationalised for the purpose of this thesis.

## 2.4. Variables

The chapter on the methodological approach of this thesis (chapter 3) highlights that factors that contribute to the outcomes of pharmaceutical benefit assessments were deduced from the empirical data rather than specifying them a priori. This means that any variables that the previous discussion of the relevant literature gives rise to were not employed for hypothesis-testing purposes, but rather heuristically as a set of variables to refer back to in the completion phase of the empirical analysis. In the final stages of the empirical analysis, the themes and factors that emerged were examined in light of the variables that arise from the policy paradigm literature in order to understand whether new and previously under-theorised factors arose. The following section is kept comparatively brief so as to avoid preempting the results of the empirical analysis that are discussed in chapters 6, 7 and 8.

The ideational and policy paradigm literature underlines the importance of ideas as variables in the policy process. However, due to the lack of a consistent definition of 'ideas', which results in a lack of a consistent operationalisation of the concept, methodological problems emerge when capturing ideas, or policy paradigms, as variables. Hall's (1993) account of the policy process remains the most sophisticated account of the central variables that are at play in the context of a wider framework of ideas, or paradigms. To recall, Hall (1993) conceptualises the policy process as a process in which goals, techniques and instrument settings of a policy are shaped by an 'interpretive framework' of ideas, the policy paradigm. If the three central variables – that is goals, techniques and instrument settings – are framed by a paradigm, then the logical inference is that they are also an expression of the dominant paradigm. Hall's (1993) central variables are also comparable to Kuhn's (1962) analysis of normal science as a process in which paradigms get formulated and articulated. That is to say, that looking at goals, techniques or instrument settings (Hall, 1993) or, in Kuhn's (1962) view, looking at rules, methods and instrumentation in normal processes provides an avenue for identifying paradigms. In other words, identifying the goals, techniques and instrument settings will prove useful for the purpose of identifying the dominant HTA paradigms and their effect on benefit assessment outcomes. Based on this theoretical assumption, I adapt Hall's and Kuhn's insights in a manner that allows me to elicit the dominant HTA paradigms in England and Germany.

As outlined in section 2.3., policy paradigms can be broad and narrow. The goals, techniques and instrument settings will serve as proxies for describing the narrow policy paradigm, that is the dominant HTA paradigm. In other words, the HTA paradigm is operationalised by describing the main features of the paradigm in terms of its goals, techniques and instrument settings. Given that the guiding assumption of the theoretical approach is that the HTA paradigms determine the outcomes of benefit assessments, it is reasonable to assume that one can capture these paradigms by reference to the variables that structure the policy process in which they operate. This assumption is consistent with Kuhn's (1962) view that normal science is *paradigms in operation* and that looking at variables within normal processes will provide a better understanding of the elements of a given paradigm.

The above remarks underline that two central tasks need to be fulfilled in order to understand the role of policy paradigms as variables in determining the outcome of pharmaceutical benefit assessments. Firstly, the broad health care and HTA paradigms in England and Germany have to be identified with the help of statutory frameworks and methodological guidelines. As outlined previously, the goals, techniques and instrument settings of HTA policy will help in the operationalisation of the HTA paradigm. Since these variables are shaped by an overarching health care paradigm, their operationalisation will also help with the identification of the principles of the overarching paradigms. Secondly, the operationalisation of the narrow HTA paradigms needs to be evaluated in the case study analyses in order to determine whether additional features can be identified that extend or alter the dominant HTA paradigm.

The transferability of Hall's (1993) framework to the empirical puzzle of this thesis might be questioned on the grounds that Hall's work pertains to policy-making processes, policy change and policy learning rather than to decision-making processes as an extension of policy-making when policy is implemented. However, despite the 'scientific-bureaucratic' (Harrison, Moran and Wood, 2002) nature of the HTA decision-making process, Hall's (1993) assertion that policy goals, techniques and instrument settings are central to understanding policy processes is applicable, especially if one takes into consideration Kuhn's (1962) work on the connection between the processes of normal science and paradigms. The thesis' theoretical contribution is in demonstrating the applicability of Hall's and Kuhn's work and showing how policy paradigm frameworks are relevant to a wider number of empirical questions than their current use suggests.

In summary, the variables that are relevant to answering this thesis' research question include policy paradigms as well as the goals, techniques and the instrument settings of policies. Since the operationalisation of policy paradigms is challenging, the goals, techniques and instrument settings – i.e. the variables of the 'normal' decision-making process - of HTA policy in England and Germany are used as proxies for the dominant HTA paradigms. The underlying intellectual premise is that the way these variables are constructed in the statutory and methodological frameworks is a reflection of ideas and values that are contained within the overarching broader paradigm that shapes the narrower paradigms in a given policy area. When the process of identifying policy paradigms is supplemented by the dominant themes that emerge in the empirical analysis, the result is an in-depth understanding of what the HTA paradigms in England and Germany are, how they are established in normal practice and how this impacts on the outcome of pharmaceutical benefit assessments. The variables serve the primary function of assisting in the operationalisation of policy paradigms rather than guiding the empirical analysis, which is characterised by a more deductive process.

The next section provides a brief introduction to the concept of evidence and rules of evidence in decision-making processes such as HTA that rely heavily on evidential information. The section further underlines how the policy paradigms of HTA are operationalised in this thesis, that is to say that Majone's (1989) concept of 'rules of evidence' is employed as a further means to identify the crucial characteristics of given HTA paradigms.

## **2.5. Rules of Evidence**

This chapter began by highlighting the important role that evidence plays in HTA processes. HTAs are built on the intellectual premise that scientific evidence should and can inform decision-making in health care (Lehoux and Blume, 2000; Van Der Wilt and Reuzel, 1998). I argue that policy paradigms, as outlined above, influence the way key questions concerning the scientific evidence in a given benefit assessment are approached. That is not to say that evidence questions are the sole variable that is influenced by policy paradigms and that shapes the outcome of benefit assessments. However, the empirical analysis undertaken in this thesis did give rise to the centrality of evidential questions in assessment processes in that five out of six themes that emerged from the data analysis are connected to questions around evidence (see table

3.5.). This, along with the fact that this thesis' research question arises from an empirical puzzle in which the evidence base, i.e. the clinical studies on a pharmaceutical product, is controlled for, necessitates a brief excursion to the discussions on the role of 'evidence' in different strands of literature in order to ascertain how controversies around 'evidence' are commonly explained. The excursion will highlight that the process of interpretation can be thought of as a process in which different 'rules of evidence' (Majone, 1989) are applied to a similar evidence base. These rules of evidence, which are similar to Kuhn's (1962, p. 40) 'rules of the game', can be explained by reference to HTA policy paradigms, as outlined in the previous sections of this chapter.

The extant literature emphasises the contextual nature of evidence, that is that the significance of evidence depends on what is relevant in a given policy or other context (e.g. Van Der Wilt and Reuzel 1998). This emphasis further strengthens the argument that policy paradigms are useful in explaining different approaches to the same evidence questions in HTAs as paradigms may provide the researcher with a means to capture the 'context' within which pieces of evidence are interpreted. Theories on the use and conceptualisation of evidence can be found in the literature on the philosophy of science, evidence-based policy-making, policy analysis and knowledge utilisation, to name but a few (e.g. Weiss, 1979; Weiss, 1999). Moreover, the role of randomised controlled trials (RCTs) as the allegedly most reliable form of evidence in HTAs and evidence-based medicine has recently been the subject of research in disciplines such as health policy, health ethics, medicine and others (e.g. Rawlins, 2012). This section does not discuss all of the approaches contained in these literatures in detail, but highlights a few of the main arguments that are raised and that are of relevance to this thesis.

Uncertainty is an inevitable part in any research endeavour due to the fact that there are still a large number of scientific and social phenomena that current knowledge cannot yet explain. How policy institutions, research bodies and professionals deal with questions of uncertainty specifically, and questions of interpretation generally, is a crucial question when making decisions based on evidence. In discussing the limits of evidence-based policies, Pawson, Wong and Owen summarise the predicament around 'evidence' in the following manner: "We seek to justify policy decisions on the basis of "known knowns". The real problem is what to make of the "known unknowns" and the even more troubling "unknown unknowns" " (Pawson, Wong and Owen, 2011, p. 519).

They go on to assert that it is almost inevitable that “[...] the knowledge base falls short of absolute, indubitable truth” (Pawson, Wong and Owen, 2011, p. 519).

Pawson, Wong and Owen (2011) exemplify their assertions by conducting a systematic review of the evidence on the efficacy of a smoking-ban policy in vehicles carrying children. In doing so, they show that most of the available evidence is correlational rather than causal and that a number of confounding factors remain unaccounted for. Therefore the evidence provides no definite answer to the above question. They conclude that: “Measurements never stand alone; they always need interpreting. There are always rival explanations of any experimental result [...]” (Pawson, Wong and Owen, 2011, p. 542).

Arguments comparable to those of Pawson, Wong and Owen (2011) are found in the literature on evidence-based medicine, HTAs and the philosophy of science. Goldenberg illustrates that critics of the view that science and evidence are infallible argue that: “[...] observations are not “givens” or “data”, but are always the product of interpretation (in light of our background assumptions)” (Goldenberg, 2006, p. 2623). Similarly, in presenting the advantages and disadvantages of using hierarchies of evidence, i.e. ranking different forms of evidence according to their alleged quality, Rawlins argues that: “[...] decision makers must exercise judgement; hierarchies are no substitute” (Rawlins, 2012, p. 1). The aforementioned viewpoints support the theoretical argument of this thesis, i.e. because HTAs are based on evidence, and because the evidence base is controlled for in England and Germany, outcomes of benefit assessment are likely to be explicable by looking at how evidence is *interpreted* in the context of diverging HTA paradigms. In order to answer this thesis research question it is therefore crucial to understand how evidence is interpreted and how one might explain why it is interpreted in a certain way.

Cartwright and Hardie (2012) and Majone (1989) speak of ‘relevance’ in order to explain differences in the interpretation of evidence. A given evidence base, the methods applied to it and the resulting interpretation must be relevant to the policy question and context one is operating in. Cartwright and Hardie (2012) offer this advice to policy-makers by arguing that evidence-based policies will only prove successful if the evidence used is relevant to the given context. They caution policy-makers to rely too heavily on RCTs because:

An RCT gives you an important true fact, that this has worked somewhere, [...]. But that won't make for a secure conclusion if the evidence is weak in support of the right kinds of answers to the other questions you need to ask (Cartwright and Hardie, 2012, p. 9)<sup>10</sup>.

Cartwright and Hardie (2012) label this a 'theory of relevance' in which facts are good, but meaningless if the relevance to the given national and local context is not ensured. They assert that facts and evidence are made relevant to a given context through the process of reasoning and argumentation in which the relevance of certain pieces of evidence over others will become clear.

Majone (1989) also highlights the crucial role of evidence as one feature in an argumentative process in policy-making. He introduces the concept of 'rules of evidence' (Majone, 1989, p. 11) – a concept used in this thesis to operationalise how different HTA bodies interpret evidence - to explain why certain pieces of evidence are deemed more relevant than others in the process of argumentation. According to Majone:

Selecting inappropriate data or models, placing them at a wrong point in the argument, or choosing a style of presentation that is not suitable for the intended audience, can destroy the effectiveness of information used as evidence, regardless of its intrinsic cognitive value. [...] evidence must be evaluated in accordance with a number of factors peculiar to a given situation, such as [...] the prevailing rules of evidence [...] (Majone, 1989, p. 11).

Majone subsumes the answers that different systems offer to challenges surrounding evidence questions under the term 'rules of evidence':

When the issues under discussion require complex patterns of reasoning and large amounts of data of doubtful reliability and relevance, explicit rules of evidence become particularly important (Majone, 1989, p. 10).

---

<sup>10</sup> A detailed account of the criticisms of RCTs is not provided in this thesis. This is because this thesis does not seek to contribute to this strand of literature, but seeks to elicit how similar pieces of evidence such as RCTs are interpreted differently in different context. Whether or not RCTs are the 'right' evidence for HTAs is irrelevant as the fact that NICE and the FJC/IQWiG regard RCTs as the highest quality evidence demands an analysis of how they deal with similar challenges contained in RCTs rather than questioning their intrinsic value. For the purpose of this chapter, suffice it to say that RCTs have been criticised for, amongst other aspects, being unrepresentative of the clinical reality in which a pharmaceutical product will be used as the strictly controlled experimental design of RCTs means that only patients with few to no co-morbidities are included in the clinical trials. In this sense, the clinical trial patient group might not always reflect the group of patients likely to receive the medication in the 'real-life' clinical context.

Rules of evidence include the criteria that are being applied in evaluating the evidence, the questions that are asked in this process, what counts as evidence in the first place and what methods are used to analyse the data (Majone, 1989). When different rules of evidence are applied to the same data, the outcome might be a different assessment of what the evidence proves or does not prove. In this sense, different rules of evidence can help explain why the outcomes of benefit assessments on the same pharmaceutical products in England and Germany are dissimilar in some cases. When examining what determines the outcome of benefit assessments, this thesis therefore employs Majone's (1989) concept of 'rules of evidence' to describe the criteria, thresholds and values that are applied to the evidence base in the assessment process. Since the rules of evidence can be found in what Kuhn (1962) would most likely label the 'normal' processes of decision-making, they are in turn a reflection of how a paradigm is formulated and established.

Majone's (1989) and Cartwright's and Hardie's (2012) elaborations on evidence as a concept that acquires meaning according to what is relevant in a process of argumentation are transferable to the research question at hand. HTA processes can be conceptualised as argumentative processes in which different interpretations are applied to a similar evidence base and 'argued out' according to what is deemed feasible and desirable within the given national context. Majone's and Cartwright's and Hardie's arguments on the importance of relevance necessitates an examination of what facts and pieces of evidence are relevant in HTA processes in England and Germany. However, while they present convincing arguments for the importance of 'relevant' evidence, their assertions fall short of offering a conceptual framework that can explain this relevance. In an effort to address this shortcoming, this thesis expands on Majone's and Cartwright's and Hardie's theories of relevance by using policy paradigms as a framework to explain what shapes the relevance of certain evidence questions over others in the first place.

In conclusion, the brief overview of how the role of evidence has been discussed in different strands of literature gives rise to two theoretical arguments that are key to addressing this thesis' research question. Firstly, evidence always requires interpretation and judgement. As Goldenberg argues: "A lesson learned from the philosophy of science is that evidence is not self-apparent or "given" when gathered from even the most idealized and controlled observational setting" (Goldenberg, 2006, p. 2630). From this perspective, even RCTs, frequently labeled the 'gold standard' of evidence-based

approaches (e.g. Milewa and Barry, 2005; Rawlins, 2012), do not give rise to infallible results.

The second theoretical argument that emerges from Cartwright and Hardie (2012) and Majone (1989) is that evidence acquires meaning by making it relevant to the given context. Even the ostensibly ‘best’ evidence is rendered meaningless if it does not answer questions that are relevant in the given policy context and if it does not meet the standards that are contained within the rules of evidence (Majone, 1989). For the purpose of this thesis, the process of interpreting the evidence can thus be conceptualised as a process in which evidence is made relevant in a given setting by applying rules to it. The rules of evidence emerge from different policy paradigms that are normalised in different countries. These policy paradigms tell decision-makers what evidence is relevant in a given national context. In this sense, this thesis links the notion of evidence with that of a paradigm. In doing so, it brings different branches of literature together in order to examine what determines pharmaceutical benefit assessments.

The next section provides a brief overview of the alternative theoretical models that were considered during the research process of this thesis as well as an explanation of why these approaches were not pursued.

## **2.6. Alternative Theoretical Approaches**

Several alternative theoretical approaches were explored in the search of the most compelling one in addressing the question of what determines pharmaceutical benefit assessments in health care systems that employ formalised HTA procedures. The alternatives that were considered include Sabatier’s (1988) advocacy coalition framework, Haas’ (1992) work on epistemic communities and, albeit to a lesser extent, principal-agent theories. The following paragraphs provide a succinct explanation for why these alternatives were ultimately not pursued.

Sabatier conceptualises public policies as belief systems:

[...] public policies/programs incorporate theories about how to achieve their objectives and thus can be conceptualized in much of the same way as belief systems. They involve value priorities [...], perceptions of world states [...] (Sabatier, 1988, pp. 131-132).

A crucial element of Sabatier's theory is the aggregation of actors in a policy subsystem into "[...] "advocacy coalitions", each composed of people from various governmental and private organizations that [...] (1) share a set of normative and causal beliefs and (2) engage in a nontrivial degree of coordinated activity over time" (Sabatier and Jenkins-Smith, 1999, p. 120). Sabatier's (1988) focus on values and 'perceptions of world states' illustrates that he credits ideational variables with much of the same importance as Hall (1989; 1993) does. However, in contrast to Hall, Sabatier (1988; 1999) argues that constructs of ideas and values are specific to a group of policy actors who share the same beliefs over time and who strive to shape policy according to these beliefs. An application of Sabatier's theories therefore demands an identification of existing advocacy coalition groups in a given policy area. Once the advocacy coalitions have been identified, Sabatier's (1988) framework suggests that policy outcomes can be explained with reference to the role these coalitions, and their belief systems, play in the policy arena one is examining.

Sabatier's (1988) approach has an appeal to it because he attributes a central role to the values and worldviews held by advocacy coalitions who shape the policy process. However, the approach goes beyond what is initially required in answering this thesis' research question. The aim of this thesis is – as a first step – to understand what factors determine the outcomes of pharmaceutical benefit assessment processes. This task demands an identification of the factors as they arise from the empirical data. Once these factors have been identified, one could analyse whether they are a reflection of the dominance of one advocacy coalition's belief system over another. However, this would be a different research endeavour from the one at hand. It would require a shift from the focus on the benefit assessment process to a focus on the stakeholders that are involved in the process in an effort to elicit their belief systems. Moreover, the advocacy coalition framework is designed to examine policy change over a long period of time in an effort to explain change rather than to examine specific processes of decision-making in what could be labeled the implementation phase of the policy cycle.

While the appeal of Sabatier's framework in the context of HTA policy is not discarded, this thesis presents the step that precedes the application of Sabatier's theories to the empirical puzzle at hand. That is to say, the HTA policy paradigms need to be identified first before examining how they may have changed over time and before exploring whether they can be associated with the belief systems of certain advocacy coalition groups. In summary, while Sabatier's theories are insightful, they go beyond

the scope of this research project and should be viewed as a model framework to be applied in future research.

Similarly to Sabatier's work, Haas' arguments on the role of epistemic communities, conceptualised as a group of knowledge-based experts (Haas, 1992, p. 3), offer valuable insights. Haas defines an epistemic community as "[...] a network of professionals with recognized expertise and competence in a particular domain and an authoritative claim to policy-relevant knowledge within that [...] issue-area" (Haas, 1992, p. 3). Haas studies the role of epistemic communities in shaping international policy coordination. In a similar fashion, one could choose to study the role of epistemic communities in the HTA policy field, especially because the field is one in which the demand for professional expertise is high. However, once again this would go beyond the scope of the research question at hand.

It is interesting that both Sabatier's and Haas' theories centre on the role of communities or groups of stakeholders as owners and advocators of specific values and belief systems. This is interesting because Kuhn's (1962) original work on paradigms also identified them as a construct that was inherent to a specific group, namely that of scientists. Due to the fact that HTAs are based on expert knowledge and technical expertise in an area of political salience, addressing different research questions through the theoretical models by Sabatier (1988) and Sabatier and Jenkins-Smith (1999) as well as Haas (1992) presents a worthwhile research endeavour in its own right. While the application of said theories is beyond the scope of this thesis, the results of the empirical analysis may provide a useful starting point for such future research as it provides an identification of the dominant HTA policy paradigms, which may or may not align with belief systems of certain advocacy coalitions or epistemic communities.

Finally, the usefulness of principal-agent approaches in explaining this thesis empirical puzzle was explored. Principal-agent theories are theories of delegation in which a principal, e.g. a policy-maker, delegates certain responsibilities to an agent, e.g. a public organisation, because he/she is lacking the technical knowledge, expertise or time to make informed decisions in a complex policy area (e.g. Guston 1996; Parsons 2005). It could be argued that NICE and the FJC as well as IQWiG are agents that have been tasked with making decisions on the benefit of pharmaceutical products to inform health care decision-making. From this perspective, pharmaceutical benefit assessment outcomes might be viewed as the agent's application of technical criteria in the decision-making process.

However, principal-agent case studies are predominantly occupied with examining how principals can ensure that their goals are pursued in a satisfactory manner by the agents (Crosson, 2013). That is to say that principals seek to minimise the uncertainty that arises from delegating tasks to other authorities (Crosson, 2013). In the context of HTA policy this approach once again offers interesting avenues for future research, but the research question at hand focuses on the specificities of how individual decisions of pharmaceutical benefit assessments are made rather than examining whether they meet the goals envisioned by policy-makers. In other words, principal-agent models would be more appropriate for examining the extent to which NICE and the FJC/IQWiG are carrying out their tasks according to the principles and goals that were stipulated by policy-makers rather than for examining what determines individual outcomes of pharmaceutical benefit assessments.

## **2.7. Conclusion**

This chapter outlined the theoretical approach that guides the research. It highlighted that paradigms are conceptualised as intellectual frameworks that determine which issues, questions and themes are considered important in the assessment process. The discussion of Kuhn's (1962) work on the connection between normal science and paradigms also highlighted that these frameworks are not static, but that they require articulation, specification and interpretation. This implies that the identification of a paradigm necessitates the study of how it gets established and formulated in 'normal' decision-making processes. Whether the focus of these processes is normal science, i.e. the activities of a research community, or decision-making in health care is irrelevant, as the main point to adopt from Kuhn (1962) is that paradigms get established and articulated through normal every-day life processes. The main criticism of the extant literature on paradigms in public policy arises from this point. Policy paradigms are currently employed as variables to explain policy change without a thorough understanding of how they are applied in the relevant normal decision-making context. However, by looking at how they are articulated one can gain insights into how fluidly or rigidly different paradigmatic elements are applied, how this impacts on outcomes and how these might in turn impact on any future change.

This chapter also emphasised that 'evidence' is a key concept that is a common feature of HTA policies in different countries. The work that emerges from the

literatures on evidence-based policy-making, policy analysis and philosophy of science suggests that the interpretation of this evidence will vary in different national contexts, depending on the factors that are considered relevant within that context. Majone's (1989) concept of the 'rules of evidence' was introduced as a means to operationalise the divergent sets of criteria that are applied to the same evidence base in different HTA systems. In an effort to capture these rules of evidence and to explain how they impact on pharmaceutical benefit assessment outcomes, this thesis employs a framework of policy paradigms that suggests that rules of evidence are both an expression of, and shaped by, narrow HTA paradigms and broader health care paradigms that operate in given HTA contexts. In doing so, it creates a link between different branches of public policy literature and explores the usefulness of extending Kuhn's (1962) work on 'normal science' to the field of HTA.

The next chapter focuses on the research design that was employed in order to address the question of what determines the outcome of pharmaceutical benefit assessments in health care systems that employ formalised HTA procedures.

## **Chapter 3**

### **Research Design**

#### **3.0. Research Aim**

The review of the literature on health technology assessment (HTA) and the political salience of the field gave rise to the following research question: What determines the outcome of pharmaceutical benefit assessments in countries that employ formalised HTA procedures? The empirical puzzle that guides this question is the variance in pharmaceutical benefit assessment outcomes across different countries that employ HTA procedures. This variance has been shown in several studies, for example in Kanavos, et al. (2010) and Sorenson (2009). However, studies on this subject fall short of providing an in-depth account of how processes and decision-making criteria impact on the final outcome in pharmaceutical benefit assessment. Moreover, previous studies are preoccupied with examining cases in which outcomes differ, thereby giving rise to the danger of selection biases that emerge when sampling cases on the dependent variable. The underlying assumption seems to be that these cases are more interesting than the cases in which outcomes are similar because similar outcomes suggest similar appraisals of a case. The purpose of this thesis is to contribute to a more in-depth understanding of how outcomes are arrived at, both in cases where the final outcome differed and in cases where the final outcome was similar. Incidentally, the use of process-tracing methods to analyse the HTA decision-making processes illustrated that cases with a similar assessment outcome offer as much insight, if not more, into what determines pharmaceutical benefit assessments as cases with different outcomes.

Against the background of this thesis' research question and the empirical puzzle, the aim of this research is to explore the issues, or the variables, that contribute to the outcome of pharmaceutical benefit assessments. In addition to the aim of understanding these variables, this research study also has an evaluative element in the sense that it seeks to evaluate the effect of HTA policy paradigms on pharmaceutical benefit assessments. In other words, the aim of understanding what variables contribute, and how, to the outcome of pharmaceutical benefit assessments is connected to the empirical contribution that this thesis strives to make and the evaluation of the effect of HTA policy paradigms reflects the theoretical contribution the thesis aspires to. It

should be noted here that the evaluative character of the research aim does not imply engaging in policy evaluation in a temporal comparative fashion, i.e. comparing the policy area post- and pre-introduction of HTA policies, but rather in a spatial comparative fashion, that is evaluating the effects of different HTA paradigms in different health policy contexts.

The literature on HTA processes gives rise to a number of variables that can be explored as part of this research, but it also suggests that the way in which outcomes are arrived at is a result of complex processes. The various dimensions, e.g. the political, methodological and ethical dimensions, of HTA procedures are complex units that this study needs to address. This study starts from the assumption that the different dimensions of HTA are not easily delimited from each other and that each is likely to contribute, to some extent, to the final outcome of the pharmaceutical benefit assessment. Guided by theories on ideas and policy paradigms that were outlined in the previous chapter, I assume that the outcome of pharmaceutical benefit assessment depends on the dominant HTA policy-making paradigm in a given country. This assumption has implications for the research design in that the research methods chosen need to be able to explain the role of complex variables that might be institutional and/or ideational in kind. They also need to cater for the possibility that a set of variables, and the paradigmatic interpretation thereof, lead to similar and different outcomes in different health care contexts. That is to say, that the mere presence or absence of specific variables in HTA decision-making might be less significant than how these variables are framed and interpreted within a given context.

### **3.1. Methodology**

In order to answer the research question a case-oriented comparative approach is employed. Several authors (e.g. della Porta, 2008; Yin, 2009; Ragin, 1994; George and Bennett, 2005) highlight the benefits of comparative and case study approaches when the research aim is to understand and explain complex processes in which one has to allow for the possibility that *how* variables matter carries equal or more weight than the question whether they are present or absent from a causal process. This explains the methodological choice of employing a qualitative case-oriented comparative approach because it allows for a contextual and interpretive analysis of the factors that determine the outcome of pharmaceutical benefit assessments.

Della Porta (2008) distinguishes between variable-oriented research that “[...] aims at establishing generalized relationships between variables, while case-oriented research seeks to understand complex units” (della Porta, 2008, p. 198). A case-oriented comparative approach thus presents a suitable methodology because it allows for the possibility that there may be different factors that contribute to the same or different outcomes in different health policy contexts. Della Porta refers to this possibility as plural causation, i.e. an effect can have different causes in different contexts (della Porta, 2008, p. 205). Considering that the outcome of pharmaceutical benefit assessments in different countries can be similar despite being based on different HTA policy paradigms, della Porta’s theory of plural causation is a valuable methodological starting point when exploring how benefit decisions are arrived at.

The idea of plural causation (della Porta, 2008) is at the heart of what the literature outlines as the benefits of comparative and case study approaches, despite authors using different terminology to describe this idea. For example, Ragin stipulates that comparative analysis:

[...] allows for the possibility that there may be several combinations of conditions that generate the same general outcome [...] [It] proceeds by comparing the configuration of causes and not by comparing presence or absence of each condition with presence and absence of the outcome (Ragin, 1994, p. 118).

Ragin’s stipulation adds to that of della Porta by underlining that causal processes can have different effects but similar causes or the same effect but different causes as a result of “combinations of conditions” (Ragin, 1994). This unmasks a complexity that needs to be considered when studying heterogeneous policy processes such as pharmaceutical benefit assessments. As such comparative approaches seek to uncover causal *patterns* rather than causes (Ragin, 1994). George and Bennett (2005) make a similar point with regards to the advantages of case study approaches when they highlight: “[...] the methods’ ability to contribute to the development of theories that can accommodate various forms of complex causality” (George and Bennett, 2005, p. 5).

The idea of “complex causality” (George and Bennett, 2005), or plural causation (della Porta, 2008), is key to understanding how I approach this thesis’ research question. Based on the extant literature, I work on the assumption that what determines

the outcome of pharmaceutical benefit assessments cannot be explained by reference to a set of easily measurable variables whose presence or absence lead to a particular outcome, but that the way ideational variables are conceptualised and operationalised contributes greatly to what determines the outcome. The idea of plural causation, or complex causality, lies at the centre of this assumption as it potentially explains why pharmaceutical benefit assessment outcomes can be different despite being based on the same pieces of evidence on, for example, clinical effectiveness. Clinical effectiveness could be used as an independent variable and measured by assigning positive or negative scores to it, depending on the results of clinical trials and the like. However, such an approach would only confirm that there are differences in outcomes, but one would have no way of explaining these outcomes, whereas the concept of complex causality allows one to examine *how* decisions are arrived at in order to explain the outcomes. Additionally, and perhaps more interestingly, employing a case-oriented comparative approach that allows for plural causation also disclosed cases in which similar outcomes were arrived at by different reasoning such as in the case of Telaprevir (table 7.2.). Such cases would have remained largely undetected when employing a methodology that does not incorporate the idea of complex causality.

In addition to the above statements on the advantages of a qualitative case-study approach, this thesis greatly benefits from the comparative component it features. There is a long tradition of comparative studies in health policy and this thesis can be viewed in the context thereof (e.g. Marmor, Freeman and Okma, 2009; Moran, 1999; Immergut, 1992; Giaimo 1995, 2001; Aaron and Schwartz, 1984). However, previous comparative studies on different aspects of health policies have focused on the comparison of institutional and structural variables to explain certain policy choices. In contrast, this study focuses on how ideational variables such as policy paradigms function in the context of different health care systems. One could have chosen to approach the question of the role of policy paradigms in determining HTA outcomes by examining one case study. However, by approaching the question with a comparative lens this study offers more room for inference and interpretation with regards to the similarities and differences that are at play when HTA paradigms are articulated. In other words, the results of the empirical analysis of pharmaceutical benefit assessment cases in one country acquire meaning when compared and contrasted to those of another country.

In practice, this research represents a small-N multiple-case study with an embedded design in that two cases (England and Germany) of HTA systems and their

paradigms are compared by way of examining smaller, embedded, units of analysis in the form of pharmaceutical products that have been assessed and appraised in both case study countries. Yin defines a case study as:

[...] an empirical inquiry that [a.] investigates a contemporary phenomenon in depth and within its real-life context, especially when [b.] the boundaries between the phenomenon and context are not clearly evident (Yin, 2009, p. 18).

This definition once again underlines the appropriateness of employing a comparative case study approach as HTA procedures are very much a contemporary phenomenon. As iterated in the section on the theoretical framework of this thesis what Yin refers to as the boundaries between phenomenon and context, i.e. the boundaries between pharmaceutical benefit assessment outcomes and institutional and ideational variables, are not clearly discernable in this research. A case study approach helps to shed light on the boundaries between variables, or on how variables interact to contribute a particular outcome.

Pharmaceutical products were chosen as the embedded units of analysis for this case study because they represent a unit in which important factors such as evidence base and disease characteristics can be controlled for. At the beginning of a HTA process the pharmaceutical products, the disease indications for which they are licensed and the evidence that is available are the same in both Germany and England. As will be outlined in chapter 4 this is because the licensing procedures for new pharmaceutical products are largely centralised in the European Union (EU) with the European Medicines Agency (EMA) in charge of granting marketing authorisations for new products. While the products are the same, they undergo different HTA decision-making processes at the Federal Joint Committee (FJC), the Institute for Quality and Efficiency in Health Care (IQWiG) and the National Institute for Health and Care Excellence (NICE), especially with regards to what evidence is deemed appropriate, why and what criteria need to be met in order to prove benefit. This thesis analyses the processes by which the same pharmaceutical products for the same disease categories are assessed in England and Germany, thus examining how the considerations in the final decision differ from each other and whether this affects the outcome of the decisions. The details on the case selection are provided in the next section of this chapter.

During the process of choosing an appropriate method for answering this thesis' research question I also considered including a quantitative analytical approach. However, due to the fact that my research aim is to understand the complex processes that give rise to pharmaceutical benefit assessment outcomes and to evaluate the role of policy paradigms in this process, I decided against the inclusion of a quantitative analysis. This is mainly because, despite the fact that I could have included more cases, a quantitative analysis would not have given me insights with regards to how and why the decisions were arrived at and what issues featured in the decision-making process. Having said that, a quantitative analysis that examines 'harder' variables such as costs and extent of clinical effectiveness is one way in which this research might be expanded upon in the future. However, as della Porta (2008) points out, variable-oriented, more quantitative approaches, offer little to no room for plural causation. This is the main reason why, for the purpose of this study, they were not considered an appropriate method for answering the research question. Finally, Dixon's and Poteliakhoff's assessment of current comparative health policy analysis as a field that is no longer based on "[...] classifications of health systems and crude rankings, but on studies that try and understand more deeply what works, where and why [...]" (Dixon and Poteliakhoff, 2012, p. 1) further supports the use of a qualitative comparative case study.

### **3.2. Case Selection**

In case study research the principal criterion to apply when selecting cases is that of relevance to the research aim (George and Bennett, 2005; della Porta, 2008). The research question of this thesis indicates that I am interested in cross-national insights on pharmaceutical benefit assessment processes in order to understand which factors determine the outcomes of these processes. The selection of cases needs to reflect this interest.

The case selection with regards to the countries that were analysed was driven by a search for countries in which pharmaceutical benefit assessment is carried out on a *national* level. The emphasis on the national level ensured that the case selection was relevant to the research question's focus on 'countries that employ formalised HTA procedure'. The emphasis also meant pharmaceutical benefit assessment processes that inform regional priorities of health care decision-making were excluded from the case selection processes as these were not considered relevant to the research question.

Additionally, the theoretical framework employed in this thesis necessitated that pharmaceutical benefit assessments are closely tied to policy-making in a given country so as to be able to analyse the effects of different policy paradigms. The alternative would have been to examine pharmaceutical benefit assessments in research settings such as universities and research foundations, but these do not usually inform policy-making and would not have provided insights on public policy paradigms. However, considering Kuhn's (1962) arguments on normal science and paradigms, a potential area for future research would be an analysis of whether the HTA policy paradigms in research settings differ from those in national settings and what effect this has on the way a product or health care intervention is assessed.

For reasons of comparability I restricted the search for appropriate cases to countries with public (tax-based) or statutory health insurance (SHI) health care systems in member states of the Organisation for Economic Co-operation and Development (OECD) because the extant literature suggests that HTA decision-making represents a popular policy for decision-makers in these countries (e.g. Sorenson, 2009). The goals and priorities of health care in non-OECD states are very different from those in OECD member states even if HTA instruments are employed. Examining this thesis' research question in the context of non-OECD member states is an interesting opportunity for future research. However, as a first step in identifying the determinants of HTA outcomes and the role of paradigms therein, my focus was on countries with a tradition of providing public health care for their citizens. Effectively, this meant that the case search was restricted to industrialised countries with established health care systems and comparable income levels. These restrictions left me with a number of possible cases to choose from.

In Europe the development of HTA as a policy-aiding instrument has developed steadily since the 1970s. The first HTA programmes and organisations were established in France and Spain, followed by Sweden and the Netherlands (Velasco-Garrido and Busse, 2005). Today HTA exists as an element of health care decision-making in most European countries, albeit in different formats and with different decision-making remits (Sorenson and Calkidou, 2012; Landwehr and Böhm, 2013). Similarly, Australia and Zealand have established HTA procedures to inform pharmaceutical coverage decisions. I examined the HTA structures in relation to pharmaceutical benefit assessments in Europe, Asia as well as Australia and New Zealand. This gave rise to a number of interesting and comparable cases that would have been relevant to this

study's research aim. For example, cases of HTA decision-making in Scandinavia and the Netherlands would have represented interesting cases as these countries have a long tradition of including HTAs in their decision-making, which is known for its emphasis on values such as equity.

In the end, a pragmatic decision was taken with regards to what was going to be feasible within the boundaries of this research project. The pragmatic decision was a result of two main considerations. Firstly, language barriers had to be considered. For example, whilst most HTA bodies in Scandinavian countries publish decision documents on their respective websites, only some documents are published in English and these are usually limited to a summary of the results of the decisions. As such they would have not given me sufficient insight into the considerations that led to the outcome of pharmaceutical benefit assessments.

Secondly, the theoretical questions guiding this research led me to look for cases that exhibit a sufficient number of institutional and historical differences from which different HTA paradigms might emerge. The case selection employed can most closely be described as a most-different system design approach. Employing a most-similar system design, i.e. comparing HTA outcomes amongst social insurance based health systems or amongst publicly financed health systems only, would have run the risk of not yielding in-depth insights into how paradigms function differently. However, based on one of the main results of this research, i.e. that different paradigms do not, as a matter of course, lead to different HTA outcomes, future research should be conducted employing a most-similar system design in order to examine whether the inverse of the above finding can be observed, that is whether a similar paradigm leads to similar HTA outcomes in different countries.

In addition to the outlined pragmatic and theoretical considerations, I factored in the size of the population covered by the given health care system and the strength of the pharmaceutical sector in a given country. The strength of the pharmaceutical sector in a country measured in, for example, indicators such as the number of jobs in the sector as well as the level of investment in research and development activities (Freeman, 2009) was an important factor to consider. This is because one can reasonably assume that a given government's economic interest in the strength of the pharmaceutical sector and the lobbying efforts by this sector are comparable in countries where the sector is similarly strong. A comparative analysis of HTA outcomes in such countries sheds light on whether the outcome of pharmaceutical benefit

assessments is overly skewed in favour or against the pharmaceutical sector. In contrast, a comparison between countries in which the pharmaceutical sector is not similarly strong would have resulted in an inability to control for this variable as a contributing factor in HTA outcomes. This, along with the feasibility and theoretical considerations outlined above, led to a final choice between the National Institute for Health and Care Excellence (NICE) in England and Wales, the Federal Joint Committee (FJC) and the Institute for Quality and Efficiency in Health Care (IQWiG) in Germany and the Haute Autorité de Santé (HAS) in France.

While the comparison of all four of the above mentioned HTA bodies would have been interesting, it would have been beyond what was feasible as part of the in-depth analysis that I was keen to pursue in this thesis. I therefore selected NICE in England as well as the FJC and IQWiG and Germany as the cases for my study. Both England and Germany have introduced pharmaceutical benefit assessments as a formal element of health care decision-making, albeit with different policy directions, thereby exhibiting a number of important institutional differences. They have both established formalised HTA bodies as quasi-public institutions with formal decision-making mandates, thereby exhibiting institutional similarities. Moreover, England and Germany represent what can be considered ‘ideal-type’ examples of health care financing and structure, the former representing a tax-based system and the latter a statutory insurance-based system (Abel-Smith, 1994). Being the birthplace for their respective health care system formats makes England and Germany the ideal cases for exploring what determines pharmaceutical benefit assessment outcomes as their systems are built on very different paradigms and yet both have introduced HTA structures in recent years. By selecting England and Germany as case studies for this small-N case study I thus also hope to contribute to an understanding of how HTA paradigms differ in tax-based and statutory insurance-based health care systems.

Additionally, England and Germany are amongst the top ten global pharmaceutical markets (IMS Health, 2013 cited in ABPI, 2013). The relative equivalence between the role played by the pharmaceutical industry in, for example, the job market means that, despite their differences in health care financing and structuring, the two countries are comparable with regards to the importance of the pharmaceutical markets to their economies. Thus, any observed differences are likely to be attributable to the health care and HTA paradigms operating in these countries rather than a variation in strength of the pharmaceutical industry.

The choice for England and Germany as cases for this study offers empirical and theoretical benefits. The empirical benefit lies in the fact that the effects of HTA policy paradigms on pharmaceutical benefit assessment outcomes are explored in ‘ideal-type’ health care settings. This in turn gives rise to the theoretical advantages in that it offers a good starting point for a study that investigates the relative role played by policy paradigms in pharmaceutical assessment process. Since this is the first study of its kind, it makes sense to start the analysis using cases that are ‘ideal-type’ cases rather than ones that might be considered exceptional or similar from the beginning. At the same time, the fact that England and Germany are ‘ideal-type’ health care systems makes them highly relevant in addressing this thesis’ research aim as the results provide avenues for further examination in other health care systems. This provides opportunities for further exploration, both by myself and by other researchers interested in HTAs and policy paradigms, in order to test whether the role of policy paradigms is similar in other cases of formalised HTA procedures.

In addition to the empirical and theoretical merits of employing England and Germany as the cases for this study, it is worth noting that Anglo-German comparisons have a long research tradition in public policy, a research tradition which this thesis hopes to contribute to. There are numerous studies in environmental policy that compare Germany’s and Britain’s responses to common problems such as the reduction of vehicle emissions, addressing the problem of acid rain and the regulation of various environmental policy areas (e.g. Boehmer-Christiansen and Skea, 1991; Boehmer-Christiansen and Weidner, 1995; Weale, O’Riordan and Kramme, 1991; Sturm and Wilks, 1997). Until it was closed in 2009 the Anglo-German Foundation commissioned and published numerous comparative studies on economic, environmental and social issues in the United Kingdom and Germany (Anglo-German Foundation, 2008), which can still be accessed online. Finally, in comparative health policy Giaimo has conducted several Anglo-German comparisons, exploring various aspects of health care reforms and relationships between health care stakeholders (e.g. Giaimo 1995; 2001; Giaimo and Manow, 1999) and Moran (1999) chose the United Kingdom and Germany as cases in his work on the health care state.

### 3.3. Selection of Embedded Case Studies

The focus on pharmaceutical benefit assessments rather than on other areas such as diagnostic products that undergo HTAs arises from the benefits that this offers in terms of controlling for issues such as evidence base and disease characteristics. The embedded units of analysis that inform this study were selected according to temporal criteria. This means that the embedded case studies were not sampled on the dependent variable (namely the outcome of pharmaceutical benefit assessments), which in turn increases the internal validity of the study by ensuring that I am not sampling on the most interesting cases that may provide insights on the role of paradigms in determining HTA outcomes. Della Porta (2008) cites this as an important criterion that ensures that cases are not selected based on observations that confirm one's theory.

In Germany, the legislation that introduced the requirement for the benefit assessment of new pharmaceutical products in the form of so-called early benefit assessments<sup>11</sup> came into practice in January 2011, thus the starting point for the search for pharmaceutical products was January 2011. In the interest of providing enough time for the analysis of the empirical cases, the endpoint for the search of the pharmaceutical products that were appraised by both NICE and the FJC/IQWiG was the end of August 2012. A search of NICE's, the FJC's and IQWiG's websites respectively showed that during this time a total number of ten of the same pharmaceutical products were appraised by the organisations. The temporal delimitation gave rise to ten products that represent the units of analysis embedded in this comparative case study.

Table 3.1. provides an overview of the products and the indications for which they are licensed. Incidentally, the case study selection for the embedded units of analysis based on temporal criteria gave rise to a sample of cases that is well balanced and does not over-represent particular disease areas. This is especially important in light of the fact that pharmaceutical benefit assessments are frequently undertaken for cancer medicines, which are usually very expensive, especially when treatments for end-of-life cancer stages are involved. However, the sample of cases analysed in this thesis is not skewed towards cancer products. The sample includes four new pharmaceutical products for cancer (Abiraterone, Cabazitaxel, Eribulin and Ipilimumab), four for chronic diseases (Boceprevir, Fingolimod, Retigabin and Telaprevir), one that is used as

---

<sup>11</sup> The Act on the Reform of the Market for Medicinal Products (Arzneimittelmarktneuordnungsgesetz – AMNOG)

a preventative measure (Apixaban) and one (Ticagrelor) that potentially affects a large number of patients as the prevalence of coronary diseases continues to increase. Considering this balanced representation of different disease characteristics and areas, there was no need to adjust for biases, which may have demanded the selection of further cases to ensure that the results of the analysis are not skewed.

**TABLE 3.1. – Overview of Pharmaceutical Products**

<b>Product Name</b>	<b>Indication</b>
1. Abiraterone	Prostate Cancer
2. Apixaban	Prevention of thromboembolic events (blood clots) after hip or knee replacements
3. Boceprevir	Chronic Hepatitis C, Genotype 1
4. Cabazitaxel	Prostate Cancer
5. Eribulin	Advanced Breast Cancer
6. Fingolimod	Multiple Sclerosis
7. Ipilimumab	Advanced Melanoma
8. Retigabin	Epilepsy
9. Telaprevir	Chronic Hepatitis C, Genotype 1
10. Ticagrelor	Acute Coronary Syndromes

### **3.4. Variables**

Since this study takes a case-oriented research approach, the themes, factors or variables that led to specific outcomes in pharmaceutical benefit assessments arose during the process of analysing the empirical data that formed the basis for this study. However, that is not to say that there were no variables that guided the analytical process. The dependent variable studied in this thesis is the outcome of pharmaceutical benefit assessment. As shown in table 3.2., the operationalisation of the dependent variable is straightforward in the sense that there are a limited number of categories it can take. In the case of NICE the outcome of a technology appraisal can be ‘recommended’, ‘optimised’, ‘not recommended’ and ‘only in research’ (NICE, 2014a). In the case of IQWiG and the FJC the outcome of the early benefit assessment is comparatively more convoluted because a given pharmaceutical product is assigned one out of six possible benefit categories (BMJV, 2011). Nevertheless, the outcome cannot be anything but one of these six benefit categories, which is why the dependent variable remains easy to operationalise.

The question of the independent, or explanatory, variables presents more challenges than that of the dependent variable. While the theoretical framework and the literature on HTA gave rise to a range of possible variables, the aim of this research required an openness about the breadth and kind of possible variables that the collected data could give rise to, especially in order to account for instances in which a “combination of conditions” contributed to a specific outcome. Rather than treating the explanatory variables that the literature gives rise to as dichotomous units that either have or do not have an impact on benefit assessment outcomes, they were utilised as a way of interpreting the study’s findings. This ensured that additional factors, which may not yet be accounted for in the literature, could be considered as meaningful in providing an explanation of what determines the outcome of pharmaceutical benefit assessments. The approach also allowed the analysis to be receptive to instances in which independent variables *combined* to produce certain outcomes.

Table 3.2. provides an outline of the variables that are expected to play a role in explaining the dependent variable. The variables arise from the discussion on the role of policy paradigms in the previous chapter. The table also includes an overview of the factors that might be indicative for certain variables and the data sources in which these indicators are likely to appear. The overview of the independent variables is an indication of the types, i.e. ideational and institutional, of variables I expect the data might give rise to. The table thus provides guidance when interpreting the results without discarding the possibility that the empirical material might give rise to additional, previously unaccounted for, explanations for the puzzle this research addresses.

**TABLE 3.2. – Overview of Dependent and Independent Variables**

Variable	Operationalisation	Data sources
<i>Dependent variable:</i> Outcome of pharmaceutical benefit assessment arising out of formalised HTA procedures <sup>12</sup>	<p><b><u>NICE</u></b> 4 possible outcomes: Recommended/not recommended/optimised/only in research</p> <p><b><u>FJC/IQWiG</u></b> 6 possible outcomes: Major, significant, marginal additional benefit (<b>positive outcome</b>) Additional benefit not quantifiable, no additional benefit, additional benefit less than that of the comparator (<b>negative outcome</b>)</p>	<ul style="list-style-type: none"> <li>- NICE’s technology appraisals (available online)</li> <li>- FJC summary documentations on early benefit assessments (available online)</li> </ul>
<i>Independent variable no. 1:</i> Ideational factors in policy paradigms of HTA decision-making (Overarching framework)	<p><i>Indicators include:</i></p> <ol style="list-style-type: none"> <li>1) Reference to norms, values, principles, ethical ideas</li> <li>2) Social and ethical values that are considered in the decision-making process</li> </ol>	<ul style="list-style-type: none"> <li>- Methods and implementation guidelines &amp; legislative documents</li> </ul>
<i>Independent variable no. 2:</i> Institutional factors in policy paradigms: Goals/purpose/techniques of HTA (Articulation of the paradigm)	<p><i>Indicators include:</i></p> <ol style="list-style-type: none"> <li>1) Political dimensions of HTA, i.e. is the purpose of HTAs to inform the in- or exclusion of services in health care benefit baskets?</li> <li>2) Policy objectives of introducing HTA</li> <li>3) Public discourse around HTA</li> <li>4) Composition of decision-making body</li> <li>4) Degree and form of stakeholder involvement</li> </ol>	<ul style="list-style-type: none"> <li>- Legislative documents</li> <li>- Stakeholder position papers</li> <li>- Media excerpts</li> <li>- Stakeholder interviews</li> <li>- NICE’s technology appraisal documents</li> <li>- FJC summaries on early benefit assessments</li> </ul>
<i>Independent variable no. 3:</i> Ideational & institutional factors of policy paradigms: Instrument settings of HTA (Articulation of the paradigm in the ‘normal’ decision-making process)	<p><i>Indicators include:</i></p> <ol style="list-style-type: none"> <li>1) ‘Rules of evidence’ (Majone 1989), i.e. what constitutes evidence?</li> <li>2) Hurdles/thresholds for a positive assessment etc.</li> <li>4) Algorithms &amp; methods for assessment</li> </ol>	<ul style="list-style-type: none"> <li>- Methodological guidelines</li> <li>- Minutes of decision-making body meetings for an overview of issues considered in the process</li> <li>- Stakeholder interviews</li> </ul>

<sup>12</sup> See chapter 4 for a more detailed description of the possible outcomes of HTAs in England and Germany.

### **3.4.1. Dependent Variable**

The fact that the dependent variable takes one of six possible forms in the case of the FJC and IQWiG<sup>13</sup>, all of which are not straightforward ‘yes’ or ‘no’ categorisations, necessitates a word on the operationalisation of the dependent variable in comparison to the format that it takes in England. For the purposes of this thesis, the first three categories of additional benefit in Germany are assigned the score of ‘positive outcome’. That is to say that major, significant and marginal additional benefit categorisations by the FJC are interpreted as overall positive outcomes in a given pharmaceutical benefit assessment procedure, despite their differences in terms of the extent of benefit. The lower three of the six benefit categories, i.e. additional benefit not quantifiable, no additional benefit and additional benefit less than that of the comparator, are assigned a ‘negative outcome’ score, as they indicate an overall negative appraisal of a product’s additional benefit. Operationalising the dependent variable in this way supports the comparison of the outcome of benefit assessments in Germany and England by evaluating whether the outcome can be interpreted as overall positive or negative.

### **3.4.2. Independent Variables**

The selection of the independent variables gave rise to a set of challenges. These challenges are connected to the theory as well as the research aim of this study. The theory of this study requires an operationalisation of independent variables in a way that sheds light on a) what the overarching HTA decision-making paradigm in a given context is and b) how this paradigm operates in practice and c) whether the paradigm’s practical application, i.e. its articulation during the normal decision-making process, deviates from its theoretical grounding. Connected to this is the study’s ambitious research aim of disclosing if and how independent variables combine to form certain outcomes in different cases. With this in mind, independent variable one reflects the need to describe the policy paradigm as laid out in legislative and methodological documents, while independent variables two and three are connected to the operationalisation of the paradigm in practice. Methodologically this means that the

---

<sup>13</sup> See chapter 4 for a detailed explanation of the benefit categories and their meaning.

examination of independent variables two and three need to be re-connected to independent variable one in order to provide insights about how complex causality works in the case of pharmaceutical benefit assessments.

The independent variables presented in table 3.2. arise out of the variables that Hall (1993) presents in his policy paradigm framework, which is explained in chapter 2. In Hall's view institutional variables such as policy goals combine with variables such as policy instruments whilst being guided by an overarching set of values that prescribes what is or is not possible under a certain paradigm. Since the aim of this research is to understand how these sets of variables combine in pharmaceutical assessment processes and in turn to evaluate the opportunities Hall's approach offers in terms of understanding these processes, the analysis of independent variable number one – the policy paradigms of HTA decision-making – forms the basis for the interpretation of the research findings.

The independent variable one is the policy paradigms of HTA decision-making in theory. It includes the values, norms and ethical basis that HTA decision-making processes are built on in a given country. Every case of a HTA system is based on normative and ethical assumptions that can be found in documents such as the legislation that introduced the system. Independent variable number one is operationalised by way of analysing the documents that lay out the context, normative assumptions and purpose of HTA decision-making. However, the HTA policy paradigm is not just restricted to the documents pertaining to HTA mechanisms, but also to the wider health care context that embodies a set of values that need to be considered. Thus, the overarching framework of HTA decision-making is a combination of the value basis of given health care system along with how this value basis is reflected in the methodological guidelines that structure HTA processes.

Independent variables number two and three, the goals and instrument settings of HTA, are the extension of number one in that their consideration allows for an assessment of how the HTA decision-making paradigm is operationalised in practice. Even though, for ease of explanation, I outline three independent variables, it is important to reiterate that the focus of this study is on examining the connection between this variables and how they determine the outcome of the dependent variable rather than to assess whether one carries more weight than the other, notwithstanding the possibility that the research findings might indicate the relative importance of one variable over another.

The operationalisation of independent variable two and three is more complex than the operationalisation of independent variable number one. This is because in the case of the HTA policy paradigms the focus is on unmasking which values and ethical norms are mentioned in legislative and methodological documents and how they are conceptualised. The goals, techniques and instrument settings of HTA are much more fluid concepts whose role can only be disclosed by examining how they featured in different cases of pharmaceutical benefit assessments. As outlined in table 3.2. their indicators can range from institutional features such as the relative influence of stakeholders or the public discourse around a certain case to the technical feature of which methods are used to assess evidence in a given case. The list of indicators presented in the column on the operationalisation of independent variables two and three is not exhaustive because the collected data might give rise to additional factors that may be considered an indicator of either of those variables.

As the next section will show, in practice the above remarks mean that the HTA policy paradigm is analysed by means of textual analysis of statutory and methodological documents which is followed by tracing the decision-making processes in the ten embedded case studies. The latter exercise is loosely guided by a search for indicators of independent variables two and three. The results of the process-tracing exercise are then compared with the results of the textual analysis of the HTA paradigms in order to a) meet the thesis's theoretical ambition of examining the effects of policy paradigms in the area of pharmaceutical benefit assessments and b) understand the combinations of conditions that lead to certain outcomes.

### 3.5. Data Sources, Collection and Analysis

The following list presents the data sources that are analysed to answer this thesis' research question:

- Consultation and decision documents on pharmaceutical benefit assessments by NICE, the FJC and IQWiG
- Statutory documents on HTA and pharmaceutical benefit assessments
- Methods and implementation guidelines for NICE, the FJC and IQWiG
- Stakeholder interviews
- A limited number of media excerpts
- A limited number of stakeholder position papers

All of the above sources of evidence are available online with the exceptions of the stakeholder interviews, which were carried out between October 2012 and June 2013.

The publicly available consultation documents from NICE include the scoping document, the pharmaceutical manufacturer's dossier, the Evidence Review Group's (ERG) assessment of the evidence, comments from stakeholders at different stages of the consultation process, any relevant draft guidance, reviews as well as the final guidance on a given product. The available data from NICE can thus be described as detailed and inclusive of a range of views and issues that played a role during a given consultation process. Moreover, NICE's final guidance on a product includes an outline of the Appraisal Committee's (AC) reasoning on different issues in a given case.

The publicly available documents from the FJC and IQWiG are detailed, but comparatively lacking in the detailed reasoning that featured in the decision-making process. The documents include the manufacturer's dossiers, IQWiG's review of the dossier as well as the written and oral statements that were given by stakeholders during the hearing process. The latter is very detailed in that the minutes of the oral hearing proceedings are included in the documents. This proved helpful in terms of tracing the argumentative processes in individual cases. However, the grounds for decision-making presented by the FJC and IQWiG are more limited in the publicly available documents. Whether this is due to a lack of transparency or an indication that the organisations adhere to the HTA paradigm in Germany in a strict fashion that makes reason-giving superfluous is explored in the empirical chapters (chapters 6-8).

In summary, the consultation documents that are available on the HTA bodies' respective websites are, despite their differences, vast and detailed and provided a useful data base for the empirical analysis. A strength of the FJC documents is that the

minutes of the oral hearings are included, a benefit that is missing in NICE's documents as these only summarise the consultation proceedings. However, this apparent shortcoming is balanced out by outlining the decision-making reasoning in individual cases, an aspect which is comparatively weaker in FJC documents. Thus, the consultation documents on pharmaceutical benefit assessments by NICE, the FJC and IQWiG provide a solid basis for exploring what determines the outcome of these assessments.

The main work of the empirical analysis of this thesis rests on the outcome of the dependent variable and an examination of the role of the outlined independent, or indeed previously unaccounted for, variables as seen in these consultation documents. The documents, which span 500-600 pages per case in both England and Germany, are the foundation for exploring the processes, both procedural and substantive, that lead to a certain outcome. However, in order to mitigate against the outlined weaknesses of the evidence, the data was triangulated by carrying out interviews with stakeholders and analysing a limited number of media excerpts and stakeholder position papers, the details of which are described in the next sections.

In addition to the consultation documents statutory documents on HTA and pharmaceutical benefit assessments as well as methods and implementation guidelines for NICE, the FJC and IQWiG are an important data source in this thesis. Their importance arises in relation to the research aim of exploring the effects of different HTA paradigms on the outcomes of pharmaceutical benefit assessments. For the purpose of this thesis, the legislative documents, methods and implementation guidelines are interpreted as a reflection of the HTA paradigm in a given health care system. The purpose, values, processes, stakeholder involvement and methodological direction presented in these documents are viewed as the embodiment of a given policy paradigm. Therefore, while the consultation documents are central to answer the question of how policy paradigms operate in normal practice and how this impacts on assessment outcomes, the legislative, methods and implementation documents play a key role in disclosing the normative basis of the paradigms. A holistic analysis of the consultation as well as legislative and methods documents is thus presumed to shed light on the theory and practice of HTA policy paradigms in England and Germany.

Included in the above list of data sources are a limited number of media excerpts and stakeholder position papers. The term 'a limited number' is used to indicate that media excerpts and position papers were not used in a systematic fashion for every

pharmaceutical product that was analysed, but only in those cases in which the consultation documents and stakeholder interviews indicated that the public discourse around an individual case might have contributed to the outcome of the pharmaceutical benefit assessment in question.

### **3.5.1. Interview Process**

The evidence collected from the consultation, statutory and methods documents was triangulated by carrying out 23 semi-structured interviews with stakeholders involved in the respective assessment and appraisal processes. The sample of interviewees arose out of the investigation of which stakeholders were involved in the individual pharmaceutical benefit assessment processes. That is to say that a purposive sampling process was used in order to identify those employees and representatives of decision-making bodies, professional physician associations, patient groups and pharmaceutical manufacturers that were involved. The method of randomly sampling interviewees was dispensed with as it was not deemed appropriate for the research aim of this study. As Tansey states, in process-tracing exercises:

the aim is not to draw a representative sample of a larger population of political actors that can be used as a basis to make generalizations about the full population, but to draw a sample that includes the most important political players who have participated in the political events being studied (Tansey, 2007, p. 765).

The interviewee sampling process started by drawing up a list of stakeholders for every case of pharmaceutical benefit assessment and dividing these stakeholders into the categories of ‘decision-making representative’, ‘professional physician association’, ‘patient group’ and ‘pharmaceutical manufacturer’. This list was easily drawn up on the basis of the respective consultation documents as these contain the details of every stakeholder group that was involved in a HTA process. Based on this list, the stakeholders were contacted via e-mail or phone between September 2012 and January 2013 in order to inquire whether they were available for an interview. Approximately 60 interview requests were sent and 23 individuals responded positively. This translates to a successful response rate of more than a third. After completion of the 23 interviews the possibility of conducting further interviews was considered. However, data

saturation was reached already in the sense that no new themes had emerged during the last five interviews that were conducted. The interview process was thus completed after 23 interviews.

Table 3.3. gives an overview of the category and the number of stakeholder groups that were interviewed in England and Germany. In order to safeguard the anonymity of the stakeholders, the table presents the category of stakeholders rather than the name of the stakeholder group that was interviewed.

**TABLE 3.3. – List and Number of Stakeholders Interviewed**

<b>Category of stakeholder</b>	<b>England</b>	<b>Germany</b>
Decision-making body	One former executive director of NICE	<ul style="list-style-type: none"> <li>• One FJC representative</li> <li>• One IQWiG representative</li> </ul>
Professional physician association	Three representatives	Three representatives
Patient groups and charities	Five representatives	None
Pharmaceutical manufacturers & pharmaceutical industry representatives	Two representatives	Seven representatives
<b>Total number of interviews</b>	<b>n=11</b>	<b>n=12</b>

Even though the total number of interviews, when compared to the number of potential stakeholder groups involved in each assessment process, indicates that not every stakeholder group was interviewed - mainly due to the response rate of a third -, table 3.4. shows that at least one interview was conducted for every pharmaceutical product. In addition, interviews with representatives from the decision-making bodies and the pharmaceutical industry were not case-specific as these stakeholders are involved in every benefit assessment process. As such the interviews are not skewed towards certain pharmaceutical products.

While table 3.3. shows a balanced distribution between the number of interviews conducted in England and in Germany, it also shows that the interviews in England are slightly skewed towards patient groups, whereas the interviews in Germany are skewed towards representatives of pharmaceutical manufacturers.

**TABLE 3.4. – Number of Interviews According to Product and Country**

<b>Pharmaceutical product</b>	<b>Number of interviews in England</b>	<b>Number of interviews in Germany</b>
Abiraterone	n=1 (patient charity organisation)	n=1 (pharmaceutical manufacturer)
Apixaban	n=2 (patient charity organisation and professional physician association)	
Boceprevir	n=2 (patient charity organisation and professional physician association)	
Cabazitaxel	n=1 (patient charity organisation)	n=1 (pharmaceutical manufacturer)
Eribulin	n=2 (patient charity organisations)	n=1 (professional physicians association)
Fingolimod	n=1 (pharmaceutical manufacturer)	n=1 (pharmaceutical manufacturer)
Ipilimumab		n=1 (professional physician association)
Retigabine	n=1 (professional physician association)	n=2 (pharmaceutical manufacturer and professional physician association)
Telaprevir	n=2 (patient charity organisation, pharmaceutical manufacturer and professional physician organisation)	n=1 (pharmaceutical manufacturer)
Ticagrelor		n=1 (pharmaceutical manufacturer)

In Germany the lack of responses from patient organisations was striking. I experienced not only a lack of responses from patient groups, but an overall small number of patient groups. Moreover, the patient groups were hard to identify. Most patient groups in Germany are self-help groups with few organisational and financial resources. These groups are committed to helping affected patients, but seem to play little to no role in pharmaceutical benefit assessment processes. This might partly be explained by the fact that patients are represented by a pooled patients' representation in the FJC. This representation is involved in the appraisal process of the FJC, albeit without the right to vote. Upon sending the official patient representation at the FJC an

interview request I was informed of an official policy that prohibits the representation to take part in PhD research. This explains why, despite my best efforts, unfortunately I was unsuccessful in securing an interview with a patient group in Germany. Interestingly, the lack of organised patient representation and the effect of this on HTA processes also emerged as a theme during the analysis of the empirical data.

The response rate from pharmaceutical manufacturers was more cautious in England than in Germany. I received a number of responses that explained that the respective pharmaceutical manufacturer was not available for an interview for reasons of confidentiality. In contrast, the response rates by representatives from the pharmaceutical manufacturers and industry in Germany were overwhelmingly positive. The kind of responses I received seemed to indicate that there is a real interest in research being undertaken on HTA processes, especially because the requirement for early benefit assessments is relatively new to the German system.

Despite the under-representation of interviews with patient groups and the over-representation of interviews with pharmaceutical manufacturers in Germany, and the vice-versa situation in England, the data collected contributes greatly to the empirical basis of this study. The aforementioned over- and under-representation means that the interview data runs the danger of being skewed in one direction or another. However, interview data is always more subjective than other forms of data as it contains viewpoints from individuals that might be biased towards certain positions. This means that interview data has to be handled with caution in any research. For the purposes of this study, when interview data is used the source of the data, i.e. whether it stems from a decision-maker or a pharmaceutical manufacturer, is mentioned in order to account for any potential biases that the data might give rise to. However, despite the aforementioned challenges and the need to account for potential biases, the interview process was a fruitful exercise in that it supported the importance of certain themes that were already visible in the consultation documents and gave rise to additional ones, especially with regards to what I discuss as auxiliary variables such as public pressure in chapter 8. Thus, using interview data to triangulate other data sources proved beneficial when it came to the research findings.

The aim of the interviews was to ascertain the stakeholders' perceptions on what determines the outcome of pharmaceutical benefit assessments, both generally and in the specific cases that they were involved in. This purpose is reflected in the format of the interview questions and the interview protocol. Although the interviews were

carried out in a semi-structured way, a set of questions loosely guided the interview process<sup>14</sup>. These questions included, but were not limited to, the following:

- How would you describe the main principles that guide decision-making at [NICE, the FJC, IQWiG]?
- What determined [NICE's, the FJC's, IQWiG's] ultimate decision in the case of [insert name of pharmaceutical product]?
- Do you feel your involvement in the process made a difference?

The interview protocol that is attached as an appendix shows that respondents were initially asked to describe their role and involvement in the assessment process as well as whether they felt that their involvement made a difference. These questions served the purpose of creating a relaxed atmosphere and establishing a good rapport between respondent and interviewer. The interview protocol differed slightly in England and Germany, which reflects specific themes that emerged during the analysis process of the consultation documents. That is to say, that the interviews were carried out after the analysis of the consultation documents was almost completed and this allowed for an exploration of some of the themes that emerged. This explains, for example, why the interview protocol includes a specific question on the role of 'patient relevance' in Germany. In this sense, the interview process permitted a further examination of the themes that characterise the English and the German HTA paradigms respectively.

Interviews were carried out in English in England and in German in Germany. Any quotations used from the stakeholder interviews in Germany were translated by myself. Depending on the interviewee's answers to different questions, follow-up questions were asked, thus allowing for the flexibility that marks semi-structured interview processes. Despite this flexibility all of the questions that are contained in the interview protocol were asked in every interview. The average length of interviews was between 45-60 minutes. With the consent of the interviewees, the interviews were voice recorded in order to transcribe them anonymously at a later date. The transcription of the 23 interviews was carried out by myself between April and September 2013.

---

<sup>14</sup> See appendix B for interview protocol.

### 3.5.2. Data Analysis

The data was analysed by employing methods of process-tracing and qualitative content analysis. According to George and Bennett: “Process-tracing offers the possibility of identifying different causal paths that lead to a similar outcome in different cases” (George and Bennett, 2005, p. 215) and to uncover previously neglected variables. They also highlight the advantages of using process-tracing methods to uncover causal patterns. They point out that the goal of process-tracing is not the generalisability of findings to all possible cases, but rather the uncovering of complex patterns in which variables might combine in order to shape a certain outcome. This suggests that process-tracing methods can be employed for the purpose of acquiring a deeper understanding of different sets of causal mechanisms. As the aim of this research is to understand not just which variables play a role in determining the outcome of pharmaceutical benefit assessments, but *how* they matter in individual cases, process-tracing methods are appropriate tools for analysing this thesis’ data.

By employing process-tracing exercises, this research demonstrates that there is more than one possible path that may lead to similar or dissimilar outcomes in the dependent variable. Which path is pursued and which outcome it leads to, is largely determined by opportunities and constraints that decision-makers are faced with within HTA policy paradigms. The process-tracing exercise, as presented in chapters 6, 7 and 8 illustrates that drawing conclusions based on the outcome of the dependent variable alone is not conducive to understanding what determines pharmaceutical benefit assessment outcomes in a meaningful way. This is substantiated by the empirical finding that benefit assessments can result in similar decision outcomes, even if different decision-making criteria are employed in the process, as seen in several cases such as the cases of Cabazitaxel (table 6.3.), Eribulin (table 6.4.) and Telaprevir (7.2.)

In terms of the practicalities of using process-tracing methods, it should be noted that I traced the *content* of the decision-making processes in the embedded cases of pharmaceutical benefit assessments rather than the procedural steps that resulted in the final outcome of the benefit assessments. The formalised nature of pharmaceutical benefit assessments processes in England and Germany, which is outlined in chapter 4, allowed me to assume that every case of benefit assessment went through the same respective procedural steps that are prescribed by the institutional settings in both countries. Therefore, I do not trace the structural processes anew in every case, but

rather focus on the content of the decision-making processes in these cases. This study thus represents a study of the decision-making processes and criteria rather than a study of institutional procedures. The reasoning behind decisions is considered the result of a process, namely the process of applying different decision-making values and criteria to the same cases, the outcome of which forms the focal point of the thesis. Process-tracing methods were used to trace the processes of *reasoning* and *argumentation* through which final decisions were arrived at. These processes of reasoning and argumentation are akin to what Kuhn (1962) believes happens when paradigms are articulated in the processes of normal science. The use of process-tracing methods is thus also an accurate reflection of the theoretical framework that guides this thesis.

In addition to using process-tracing methods, the consultation documents and interview transcripts as well as statutory guidelines and position papers were analysed by employing qualitative methods of content analysis. The documents were analysed with a view to uncover themes and views that play a role in determining the final outcome of pharmaceutical benefit assessments. These themes were identified during an analytical exercise in which the data was read repeatedly and marked according to the themes that emerged. The themes were then compared between each other in order to assess whether congruencies, discrepancies or additional subjects arose in different documents. Apart from one theme, which is discussed in chapter 8, the findings that resulted were largely congruent between different data sets, which strengthens the internal validity of the research findings.

The data analysis gave rise to six themes that repeatedly emerged during the argumentative processes of pharmaceutical benefit assessment in England and Germany. Table 3.5. provides an introductory overview of these themes. They are introduced in further detail in chapter 6 and then discussed in chapters 6, 7 and 8. For now, suffice it to emphasise that apart from theme six, the themes that played a role in the argumentative processes of HTAs were similar in both countries. However, the questions that were asked within these broadly similar themes differed in England and Germany. This presents the starting point for the analysis of how paradigms can explain the differential emphasis on broadly similar scientific, procedural and political themes. Due to the breadth of the collected empirical material, excerpts from the consultation documents, the interview transcripts as well as other data sources are used illustratively in the discussion of the results of the data analysis throughout the empirical chapters.

### 3.6. Concluding Remarks

This chapter provided an overview of the research design employed in addressing the research question. The outlined research methods arise from the research aim of this study, that is to gain a deeper understanding of how variables combine to determine the final outcome of pharmaceutical benefit assessments in countries that employ HTA processes. The research design is also a logical extension of the thesis' theoretical premise, which suggests that ideas and values in the form of policy paradigms play a role in determining the outcome of pharmaceutical benefit assessments. The ideas and values that are outlined as policy paradigms in chapter 5 can be captured not just by analysing which variables matter, but by tracing *how* they matter, *why* they matter, and how they are framed and understood in different HTA systems. This will provide insight into how a paradigm is established and operationalised in 'normal' decision-making processes.

The next chapter offers a descriptive overview of the institutional arrangements of health care, pharmaceutical policy and pharmaceutical benefit assessment in England and Germany. It sets the contextual scene for the analytical chapters that follow.

**TABLE 3.5. – Themes that Emerged from the Data Analysis (Ranked in order of Prevalence)**

<b>Themes</b>	<b>Properties/questions of the theme in England</b>	<b>Properties/questions of the theme in Germany</b>
<i>1. Permissibility, quality and validity of evidence</i>	<ul style="list-style-type: none"> <li>• What is accepted as evidence?</li> <li>• Evidence applicable to UK clinical practice?</li> <li>• What does the evidence say about the product?</li> <li>• Does it meet the decision-making criteria?</li> </ul>	<ul style="list-style-type: none"> <li>• What is accepted as evidence?</li> <li>• Does the available evidence reflect marketing authorisation and clinical guidelines?</li> <li>• What does the evidence say about the product?</li> <li>• Are the presented clinical endpoints patient relevant?</li> </ul>
<i>2. Choice of comparator product</i>	<ul style="list-style-type: none"> <li>• Is the choice of comparator reflective of UK clinical practice?</li> </ul>	<ul style="list-style-type: none"> <li>• Is the choice of comparator reflective of the current standard alternative?</li> </ul>
<i>3. Patient population subgroup divisions</i>	<ul style="list-style-type: none"> <li>• Do the subgroup divisions adequately reflect the patient groups likely to receive the treatment in routine clinical practice?</li> </ul>	<ul style="list-style-type: none"> <li>• Do the subgroup division adequately reflect the patient population(s) for which the product is licensed?</li> </ul>
<i>4. Operationalisation of criteria for HTA decision-making and role of algorithms</i>	<ul style="list-style-type: none"> <li>• Algorithm is applied: Cost effectiveness threshold in terms of incremental cost effectiveness ratio (ICER) per quality-adjusted life year (QALY)</li> </ul>	<ul style="list-style-type: none"> <li>• Is an algorithm for the categorisation of added benefit applied by IQWiG in assessment process?</li> <li>• Algorithm applied by the FJC not transparent (<b>stakeholder opinion</b>).</li> </ul>
<i>5. Suitability of paradigms for cases such as chronic diseases</i>	<ul style="list-style-type: none"> <li>• How to make decisions in cases where natural progression of the disease is uncertain and patients live with the illness for a long time?</li> </ul>	<ul style="list-style-type: none"> <li>• Questions around the applicability of patient relevant endpoints such as mortality in disease indications where natural progression of the disease is uncertain and patients live with the illness for a long time (<b>stakeholder opinion</b>)</li> </ul>
<i>6. The question of political power and influence: Public pressure and distribution of bargaining power of stakeholders</i>	<ul style="list-style-type: none"> <li>• What is the effect of public pressure in the form of media and patient campaigns on the final result of HTAs?</li> </ul>	<ul style="list-style-type: none"> <li>• Does the differential distribution of bargaining power of stakeholders affect the final result of HTAs?</li> </ul>

## Chapter 4

### Health Care, Pharmaceutical Policy and Pharmaceutical Benefit Assessment in England and Germany

#### 4.0. Introduction

This chapter provides an overview of the structure and financing of the health care systems in Germany and England. It introduces the reader to the features of pharmaceutical regulation and pharmaceutical benefit assessment in the two countries, including the role played by the institutions that are the focus of this study, namely the National Institute for Health and Care Excellence (NICE) in England, the Federal Joint Committee (FJC) and the Institute for Quality and Efficiency in Health Care (IQWiG) in Germany. The chapter presents an introduction to the institutional specificities that need to be considered when engaging in a comparative analysis of what determines the outcome of pharmaceutical benefit assessments based on health technology assessment (HTA) procedures.

The institutional specificities include procedural and substantive similarities and differences in the way HTA processes are conducted in Germany and England. A descriptive overview of these, including the form that the dependent variable, i.e. the outcome of pharmaceutical benefit assessments, can take in the two countries is followed by a comparative summary in section 4.9. in order highlight the first important distinction between the German and the English HTA policy paradigm. That is, while the German HTA paradigm operates on the idea that there are thresholds for the ‘right’ price of a pharmaceutical product, the English HTA paradigm operates on the idea that there are thresholds for the inclusion or exclusion of a pharmaceutical product, i.e. thresholds above which a product will no longer be considered good value for money. The concept of thresholds is thus employed for different purposes in England and in Germany, which inevitably leads to differences in how the concept is operationalised in normal practice. Whether or not this difference can explain the differences or similarities in HTA outcomes is explored in the later parts of this thesis. What follows is an overview of the health care systems, pharmaceutical policy and the HTA processes in England and Germany in order to get a sense for the institutional context in which the emerging dominant HTA paradigms are established.

England and Germany represent two different types of health care systems that are characterised as ideal types in the health policy literature. The literature distinguishes between types of financing, service provision and mode of governance in health care. Abel-Smith (1994) outlines that health care financing can take the form of public or private health expenditure or compulsory health insurance. These types of health care organisation are exemplified by the German and the English health care systems. Table 4.1. provides an overview of the main features of the health care systems. In light of this thesis' research question only the institutional and structural elements that are considered relevant to the question are presented in more detail in the following sections.

**TABLE 4.1. – Health Care in England and Germany**

<b>Characteristics</b>	<b>England</b>	<b>Germany</b>
<i>Financing</i>	Publicly financed based on taxation	Publicly financed based on statutory health insurance (SHI)
<i>Health care expenditure as % of GDP in 2012 (OECD, 2014) (figures for more recent years not available)</i>	9.3% (figure for United Kingdom)	11.3%
<i>Population covered by public health care</i>	63.2 million legal residents (Department of Health, 2013)	90% of population covered by SHI, the rest is covered under private insurance (BMG, 2013)
<i>Health care service structure</i>	Division between primary and secondary care. General practitioners (GPs) hold role as gatekeepers in the systems.	Division between ambulatory and hospital care. Ambulatory care includes patient access to specialist consultants. No strong gatekeeping role for GPs.
<i>Service charges for patients</i>	Free at the point of access. Cost-sharing for prescription charges with exemptions for certain population groups.	Free at the point of access. Cost-sharing for prescription charges and for some preventive services.
<i>Decision-making structure/Mode of governance</i>	Purchaser/provider split. Local clinical commissioning groups (CCGs) decide which services are commissioned and service level agreements are drawn up with providers.	Emphasis on self-governance. Self-governing sickness insurance funds negotiate terms, conditions and prices for health care services with providers.
<i>National/federal bodies with a decision-making mandate on health care service provision</i>	NICE's technology appraisal guidance has to be implemented by local CCGs.	FJC makes federal decisions on the minimum of services to be covered by sickness insurance funds.

## **4.1. Health Care in England**

### **4.1.1. Financing**

The health care system in England represents a publicly financed system based on taxation. It is frequently referred to as the ‘Beveridge’ model of health care financing, named after a British economist, William Beveridge, who in 1942 produced the so-called Beveridge Report “[...] which advocated free and universal health services [...]” (Aaron and Schwartz, 1984, p. 13). What was then envisaged in theory was implemented in practice in 1948 with the creation of the National Health System (NHS). Since then the NHS is financed predominantly through taxes and has a fixed annual budget to cover the services it provides. While there are other sources of income such as prescription charges paid by patients, in comparative terms the NHS in England represents a relatively ‘pure’ case of a tax-based health care system.

An annual fixed health care budget that consists of taxpayers’ money comes with its own unique challenges with regards to planning services and being accountable in the eyes of the taxpayer. The Department of Health (DH) summarises this in the following way: “The challenge faced by the NHS is how to spend that budget in a way that results in the best possible outcomes for individual patients and delivers value for money for the public” (Department of Health, 2013, p. 6).

### **4.1.2. Principles, Structure and Health Care Provision**

The most important and long-standing principle of the NHS is that its services are free at the point of access. Additional principles of the NHS are laid out in the so-called NHS Constitution (Department of Health, 2013a). The Constitution lays out seven principles that include the provision of a comprehensive service to all, access to health care services irrespective of ability to pay, the centrality of the patient, value for money and accountability to the public (Department of Health, 2013a).

With its financing based on taxation, the role played by the state is traditionally bigger in systems such as the NHS than in statutory health insurance (SHI) systems. While the government delegates responsibility for the provision and planning of health care services to autonomous actors in SHI systems, publicly financed systems mean that the responsibility, and hence accountability, of service provision lies more directly with

state organisations. Thus, while actors such as GPs are autonomous in England in that they run their own practices and purchase services for their patients from service providers, the NHS system exhibits a higher degree of centralisation in decision-making than that which we observe in health care systems with a different financing mechanism.

Most decisions on which services to fund and provide are made on a local and regional level. The reasoning behind this is that the local demographics of a region determine the health care needs of the area (Department of Health, 2013). Making decisions on a local level allows for prioritising services that are perceived to be needed the most and from which the local population benefits the most (Department of Health, 2013). Since April 2013 the regional organisations in charge of making decisions about which services to provide are the Clinical Commissioning Groups (CCGs) (Department of Health, 2013). These groups took over from the so-called Primary Care Trusts (PCTs). At a national level the NHS Commissioning Board, also known as the NHS England, make commissioning decisions that relate to rare diseases, offender health care and health care for the armed services as these may entail special requirements that need to be addressed (Department of Health, 2013, p. 7).

The allocation of health resources in England has not been without its challenges. For example, concerns about the effects of health resource allocation and priority setting have included concerns about the so-called “postcode lottery” (Ham, 2009, p. 128) of service provision in which access to treatment for a patient might depend on the area where he/she lives, thus giving rise to concerns over a regionally determined inequality of access. In 1999 these concerns led to the establishment of the National Institute for Health and Clinical Excellence (Ham, 2009), now the National Institute for Health and *Care* Excellence (NICE). By evaluating treatments and pharmaceuticals and issuing guidelines and recommendations in these areas, NICE seeks “[...] to reduce variation in the availability and quality of NHS treatments and care” (NICE, 2013). An overview of the structure and mandate of NICE is provided in a later section of this chapter. For now I briefly turn to the role of cost containment in the NHS.

### **4.1.3. Cost Containment**

In comparison to Germany, England has a long tradition of cost containment in the form of rationing or prioritisation due to its dependence on annual fixed health care budgets (Robert, 2003). Ham highlights that until the 1970s rationing occurred rather implicitly in the form of long waiting lists or denial of access to specialist services (Ham, 2009, p. 127). From the 1970s policy-makers began to acknowledge a need for putting priority setting on the health policy agenda, albeit with a reluctance to take responsibility for making decisions on which services to fund (Ham, 2009). Ham asserts that: “The reluctance of governments to take a lead in setting priorities derives from the political costs involved in taking unpopular decisions” (Ham, 2009, p. 127). This explains why, for the most part, rationing and health priority setting in England is considered an issue to be resolved at the local level through commissioning responsibilities which in turn gives rise to the previously described phenomenon of the ‘postcode lottery’ where access to services depends on where one lives (Ham, 2009).

### **4.1.4. Summary**

In contrast to Germany’s health care system, decision-makers and providers in England are constrained by annual fixed budgets that leave little room for variable spending increases. Unexpected or new developments such as a particularly severe flu season or positive recommendations by NICE which CCGs are under an obligation to fund put strains on local budgets which are not necessarily taken into account when budget allocation occurs. This means that local health care decision-makers in England are frequently faced with the question of which services can justifiably be displaced by the coverage of new services. In the past, this has given rise to regional differences in health care service provision, which policy-makers have sought to address through the introduction of NICE (Ham, 2009).

For the purposes of this thesis the analysis of the pharmaceutical benefit assessment process in England has to be undertaken against the contextual backdrop of a) an institutionally setting in which decisions on health care provision are predominantly taken at the local level by way of health care commissioning; the legal obligations of local health care commissioners include the obligation to prioritise services such as described by the government and to fund treatments and

pharmaceuticals that are recommended by NICE and b) a financing mechanism that means that local health care commissioners are constrained by annual budgets that are fixed by central government. While commissioners are payers for health care services in that they purchase them from providers, they are not payers in the sense that they decide how much money is available to them. This decision is made centrally in annual budget negotiations. This element of health care decision-making presents an important distinction to the German case in which payers hold a more active negotiating role when it comes to the level of reimbursement for services.

## **4.2. Health Care in Germany**

### **4.2.1. Financing**

The health care system in Germany represents a system that is publicly financed through compulsory social insurance. It is a statutory health insurance (SHI) system that dates back to 1883 when the Chancellor of the German Empire, Otto von Bismarck, introduced laws that mandated compulsory health insurance for industrial workers and workers of other trades (Simon, 2010). For this reason the German health care system is frequently referred to as the ‘Bismarck’ model of health care financing (Simon, 2010).

There is a statutory obligation for all residents in Germany to have health insurance which, depending on one’s salary, is provided by quasi-public, not-for-profit insurance bodies called statutory sickness funds or by private insurance bodies (Simon, 2010). The insurance contribution rates are divided between employers who currently pay a rate of 7.3% and employees who pay a rate of 8.2% of the gross salary towards statutory health insurance (BMG, 2013a). This means that in addition to monthly taxes, 15.5% of a given monthly gross salary will be deducted to pay for statutory health insurance (BMG, 2013a). There are special provisions for children, unemployed family members and unemployed members of society. Children and unemployed family members are insured free of charge with the employed members of the family and the insurance fees for citizens on social benefits are covered by the state (BMG, 2013b).

Since 2009 the SHI contributions are pooled in the so-called ‘Gesundheitsfonds’, a health fund in which SHI contributions are collated, tax-supplemented and then redistributed to sickness funds in a risk-adjusted manner (BMG, 2013c). The creation of this health fund included the introduction of the previously

mentioned uniform SHI rate of 15.5% in order to ensure planning stability for employers and employees. The introduction of a tax-based element in form of tax supplements as part of the health fund to mitigate some of the expenditure increases in health care represents a novel policy in health care financing in Germany in that it is no longer covered solely by employers and employees but supplemented directly by the state. This is an example of what Saltman (2012) means when he asserts that the traditional boundaries between health care systems, such as the distinction between tax-based and social insurance-based systems, are beginning to fade, thereby opening possibilities for a wider array of cross-national research.

#### **4.2.2. Principles, Structure and Health Care Provision**

The health care structure in Germany rests on several important principles. According to Simon (2010) they include the constitutional duty of the state to provide social services, the principle of subsidiarity, the principle of solidarity and the principle of self-administration. For the purpose of this thesis, only the principles of solidarity and self-administration are explained briefly in the following sections as they represent institutional and normative features that are key to understanding the discussion of the HTA paradigms which is presented in the next chapter.

The principle of solidarity is anchored in §1 and § 3 of the Social Code Book V (SGB V) (BMJV, 2013), the statutory framework that guides health care in Germany. This principle forms the normative pillar of the German health care system in that it outlines that health care financing is a communal responsibility. Members of the community show solidarity by paying for statutory health insurance and accepting that the fees they pay will be used when other members of the society require health care (Simon, 2010). The principle of solidarity is based on an understanding of reciprocity, which renders it acceptable because members of the SHI community know that they too have a right to access health care when they need it. As such the system represents a solidary agreement between the healthy and the ill members of society.

The principle of self-administration, or self-governance, is an institutional feature commonly found in SHI systems and it is connected to the structure of decision-making and health care provision. SHI systems are characterised by de-centralised decision-making procedures in that the state lays out the legislative mandate for health care, but delegates the responsibility of implementation to independent health care

bodies and actors (e.g. Giaimo, 1995). In Germany this delegation is manifested in the principle of self-governance. This means that self-governed sickness funds, professional physician representations, hospital organisations and professional dentist bodies negotiate the terms and conditions of service provision, remuneration scales and pricing amongst each other (Simon, 2010). It is for this reason that the German system is frequently referred to as a corporatist system which “[...] hands over certain legally defined rights of the state to self-governing institutions” (Pfaff, 2009, p. 104). In practice this means that health care decision-making in Germany is diverse and complex in that it takes place at different levels of federal and regional tiers. If an agreement cannot be reached between negotiating parties, then the matter is referred to independent bodies of arbitration which aim to solve the dispute and prevent participating parties from taking legal action.

At the federal level, decisions on which services to include or exclude from the health care benefit basket are made by the main decision-making body, the Federal Joint Committee (FJC). The next section briefly outlines the set-up of the FJC and its mandate with regards to health care decision-making. Its role in pharmaceutical benefit assessment is discussed in more detail in section 4.8..

#### **4.2.3. The Federal Joint Committee (FJC)**

The FJC is the main decision-making body in health care in Germany and as such it constitutes the most important body of the self-governing structure (Simon, 2010, p. 102). The FJC consists of five representatives of the payer (the statutory sickness funds) and provider associations (physicians, dentists and hospitals) respectively, three impartial members as well as five patient representatives (FJC, 2013). The latter may take part in deliberations and propose resolutions, but they do not have voting rights. FJC decisions and guidelines cover services such as pharmaceutical coverage, needs planning, disease management programmes (DMPs) for chronically ill patients, therapeutic and diagnostic devices. The FJC’s decisions are binding for statutory sickness funds, health care providers and SH-insured members alike (Simon, 2010, p. 102).

The division between ambulatory and hospital care in Germany is visible in the decision-making remit and regulations of the FJC. In the ambulatory care setting a new diagnostic or therapeutic method will only be reimbursed if the service has received a

positive decision from the FJC (§ 135 Abs. 1 SGB V). By contrast, in the hospital care setting a new diagnostic or therapeutic method will be reimbursed unless it has received a negative decision by the FJC (§137c SGB V). Thus, in terms of service coverage, the default position of what is provided is different in ambulatory and in hospital care. In hospital care the default position is to ‘do’ unless a service is specifically excluded, whereas in ambulatory care the default position is ‘do not do’ unless it is specifically included. Special rules pertain to the coverage of new pharmaceuticals, which are explained in section 4.8..

#### **4.2.4. Cost Containment**

In SHI systems the health care budget does not compete for funds with other tax-financed policy areas such as education or defence. However, this does not mean that SHI systems are immune to the impact from economic downturns. The fact that the SHI contribution rates are dependent on salaries indicates that the development of health care expenditure is closely linked to the economic climate. That is to say that contributions decrease when unemployment increases or salaries stagnate. This suggests that health care policy-makers in Germany are facing similar challenges as those in other countries.

To address the pressures on the German health care system that stem from rising expenditures policy-makers have introduced several reform measures. Busse (1999) discusses the reform, rationing and priority setting measures in different health care sectors in Germany. Measures have included the introduction of diagnostic-related groups (DRGs) for the payment of hospital services, various supply- and demand-side measures in the pharmaceutical policy area and “[...] maintaining stability in the average SHI contribution rate [...]” (Busse, Schreyoegg and Henke, 2005, p. 329). In 2010, policy-makers introduced a law to help curtail the prices of new pharmaceuticals in Germany, the Act on the Reform of the Market for Medicinal Products (Arzneimittelmarktneuordnungsgesetz – AMNOG) (Bundesgesetzblatt, 2010), which is described in more detail in section 4.4..

Despite these reform measures the question of whether prioritisation and rationing is happening in the German health care context is a contested issue. Calls for an open debate on prioritisation have predominantly been voiced by professional physician bodies such as ethics commission of the German Medical Association

(Zentrale Ethikkommission bei der Bundesärztekammer, 2007). These calls have not received much response from policy-makers who remain reluctant to engage in debates about rationing and prioritisation (Ärzteblatt, 2011). This reluctance was confirmed by a qualitative study by Klingler, et al. in which “respondents attested a traditional reservation towards associating health benefits with financial values” (Klingler, et al., 2013, p. 275).

The lack of a public discourse around issues of priority setting in health care is an important institutional feature to bear in mind when deconstructing the HTA paradigm in Germany. However, despite the lack of a public discourse on such questions, there are suggestions that rationing is, at least implicitly, already happening in everyday clinical settings in Germany, especially in the hospital sector (e.g. Strech, et al., 2009; Marckmann, 2009). While the extent of this implicit rationing is still unknown (Marckmann, 2010), it is predominantly attributed to reforms such as the introduction of DRGs as a method for the payment of hospital services (Altenstetter and Busse, 2005). Yet, despite this there are also reports of a surplus of funds in the health fund (F.A.Z., 2013). In August 2013 this surplus amounted to 29 billion Euros (F.A.Z., 2013), a figure that might contribute to the perceived lack of need to engage in discussion about prioritisations in health care.

#### **4.2.5. Summary**

The German health care system is characterised by a corporatist mode of governance in which autonomous, self-governing actors negotiate the terms and conditions of health care provision amongst themselves with the state mandating the minimum requirements and criteria of this provision (Altenstetter and Busse, 2005). In relation to its financing, it is dependent on the stability of SHI contribution rates, a stability that can only be achieved in favourable economic climates.

Based on the above remarks, an analysis of the early benefit assessment of pharmaceuticals has to take into consideration a) an institutionally complex bargaining system that includes a diverse set of actors, both from the payer and from the provider sides, who negotiate health care service provision and b) a financing mechanism that allows for a certain degree of flexible health care decision-making and that is, for the moment, running a surplus, but that does not shield from negative external shocks. Moreover, policy-makers and decision-makers alike appear to be reluctant to engage in

debates on priority setting and rationing. Addressing the question whether this is a genuine reluctance or a perceived one, or whether it merely reflects a lack of need to prioritise in a system that frequently runs surpluses, is beyond the scope of this thesis. However, the reluctance, be it real or perceived, might contribute to the interpretation of the results of the empirical analysis presented in chapters 6, 7 and 8.

### **4.3. Health Care in England and Germany: Concluding Remarks**

The above overview of the English and German health care system provided an introduction to the institutional characteristics that are important to consider when examining the HTA policy paradigms in both countries. At its core, this thesis presents a juxtaposition of the German health care system which is characterised by decision-making structures that are based on bargaining vis à vis the English system that is characterised by localised decision-making processes that are guided by legal obligations and annual fixed budgets subscribed by central government. I compare and contrast a system that is currently running a surplus of health care funds with one that has experienced the constraints of fixed budgets since its inception in 1948 and has thus had to face the difficult choices of rationing and prioritisation from an early point in its life cycle.

While the institutional differences are stark, the next section illustrates that they are by no means insurmountable or indeed a barrier to institutional comparison. This is because one area, namely pharmaceutical policy, has played a large role in reform efforts in both Germany and England. Reform measures in this field have been comparable, especially when it comes to the creation of so-called HTA bodies. In the next section I give a brief overview of pharmaceutical regulation, policy and pricing in both countries before outlining the workings of the main decision-making bodies in the field, NICE in England and the FJC and IQWiG in Germany.

#### 4.4. Pharmaceutical Spending, Policy and Regulation

Pharmaceutical regulation and policy includes instruments used by national governments to ensure the safety of available medicines, the responsible and efficient use and prescription thereof and measures to contain the costs of pharmaceutical spending (Seiter, 2010). It also relates to regulatory measures affecting the pharmaceutical industry, an industry which is a significant economic player in both Germany and the United Kingdom (UK) in terms of exports and job provision (Mossialos and Oliver, 2005).

As briefly outlined in chapter 1, pharmaceutical spending has been a prominent target of health care reforms in OECD countries. England and Germany are no exception to this trend. Pharmaceutical spending constitutes a large part of overall health care expenditure in the two countries. In 2011<sup>15</sup> the UK<sup>16</sup> spent approximately £13 billion on its medicines bill, that is 9.8% of the total NHS expenditure (Hawe and Cockcroft, 2013, pp. 137-138). In 2013 the sickness insurance funds in Germany spent €30 billion on medicines, that is 14.4% of their total expenditure (Statista, 2014). While these proportions of total health care spending on pharmaceuticals may not immediately appear significant at between 10-15%, their significance becomes salient when considering that pharmaceutical spending is “[...] the third biggest spending component after inpatient and outpatient care” (OECD, 2012). This explains why pharmaceutical spending has been subject to reform measures such as the introduction of patient co-payments, reference pricing, profit controls and generic drug prescriptions (Freeman, 2009) and, more recently, health technology assessments.

The expenditure on medicines in England and Germany is not the only aspect that policy-makers in both countries are concerned with in pharmaceutical policy. England and Germany are amongst the top ten global pharmaceutical markets (IMS Health cited in ABPI, 2013). In the area of research and development (R&D), for example, the pharmaceutical sector in the UK constitutes the business sector with the highest R&D expenditure. “The expenditure accounts for 28% of the total expenditure on R&D performed in UK businesses in 2011” (ONS, 2012, p. 4); this figure compares with, for example, the aerospace industry’s R&D expenditure of just 8%, which is considered the largest aerospace business in Europe (ONS, 2012). Using a slightly

---

<sup>15</sup> Figures for years later than 2011 not yet available

<sup>16</sup> Figures only available for UK and not for England alone

different example for Germany, Germany's pharmaceutical sector employs the most employees compared with other pharmaceutical sectors in Europe, making it the biggest pharmaceutical market in Europe in terms of people employed (IW and vfa, 2013).

The significant strength of the pharmaceutical sector in England and Germany adds to the complexity of pharmaceutical policy and regulation. It underlines the political and economic salience of the introduction of HTA policies because the need for containing pharmaceutical spending has to be balanced against the need to provide an attractive environment for investments in jobs and R&D efforts. The complexities of meeting competing demands in the area of pharmaceutical policy-making, which have been highlighted by the illustrative figures above, form the background against which the following sections and chapters have to be understood.

#### **4.4.1. Pharmaceutical Licensing**

Before a pharmaceutical product can be marketed in a given country, the pharmaceutical manufacturer has to acquire a license, a so-called market authorisation. In the European Union (EU) this licensing procedure has been harmonised. Most pharmaceutical products that are new to the European market go through what is called the centralised procedure (CP) which allows pharmaceutical manufacturers to apply for a market authorisation with the European Medicines Agency (EMA) (EMA, 2013). If the EMA is satisfied that the product meets the necessary standards of quality, safety and efficacy, the European Commission (EC) grants the license (EMA, 2013). This means that the pharmaceutical product in question will be licensed in all EU member states at the same time. The advantage of this procedure is that pharmaceutical manufacturers only have to go through a licensing process once and that they can use one evidence base, usually in the form of randomised controlled trials (RCTs), to acquire their license. Alternatively, manufacturers can choose to follow the decentralised procedure (DCP), thus applying for a license via national authorisation systems that are in place in every EU member state. If successful in securing a license the manufacturer may then apply for the license to be recognised in more than one member state under the so-called Mutual Recognition Procedure (EMA, 2013).

Pharmaceutical licensing in Germany was introduced in 1968 with the Medicines Act 1968, as a result of scandals surrounding the use of unsafe medicines. The agency responsible for the licensing and registration of pharmaceuticals and

medical devices is the Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM), the Federal Institute for Drugs and Medical Devices. The institute carries out its responsibilities under the Arzneimittelgesetz (AMG), the German Medicines Act. Pharmaceutical manufacturers can apply for a license directly to the BfArM under the decentralised procedure that is available to manufacturers in EU member states. As an authorisation agency, the BfArM “[...] reviews the proof of efficacy, safety and adequate pharmaceutical quality of the finished medicinal products” (BfArM, 2013) and “the pharmaceutical companies must provide proof of the pharmaceutical quality, efficacy, and safety of the product. If they fail to do so the BfArM refuses authorisation” (BfArM, 2013). The institute monitors existing and new drugs for potential side effects even after they have entered the market, as such its role does not stop when the license is granted.

The British equivalent of the BfArM is the Medicines and Healthcare Products Regulatory Agency (MHRA). The application and authorisation procedure functions in much of the same way as that of the BfArM with efficacy, quality and safety being the main criteria that have to be met. In order to assess the safety and efficacy of the product in question the pharmaceutical company supplies the MHRA with all necessary data resulting from clinical trials, toxicological tests and the like (MHRA, 2013). Monitoring of the products does not stop after licensing and the MHRA can withdraw licenses at any time (MHRA, 2013).

#### **4.4.2. Pricing and Reimbursement**

Once a medicine has been licensed decisions have to be made about its price. Even though the price setting mechanisms in England and Germany differ, they have in common that restrictions exist on the extent to which pharmaceutical manufacturers are allowed to set prices freely.

In England the prices for licensed, branded, patented pharmaceuticals are regulated via the Pharmaceutical Price Regulation Scheme (PPRS), a voluntary agreement between the Department of Health and the pharmaceutical industry represented by the Association of the British Pharmaceutical Industry (ABPI) (Department of Health, 2009). Mossialos, Walley and Mrazek refer to this form of price regulation as an “indirect price control through profit or rate-of-return regulation” (Mossialos, Walley and Mrazek, 2004, p. 11). This kind of regulation aims to regulate

manufacturers' profits. The negotiating parties negotiate profit margins on the manufacturers' return on capital (ROC) employed (Walley, Mrazek and Mossialos, 2005, p. 393). If a manufacturer exceeds this margin, it has to pay a rebate to the NHS or if it underperforms it is allowed to raise prices (Walley, Mrazek and Mossialos, 2005). Walley, Mrazek and Mossialos (2005) argue that this sort of regulation creates a stable business environment that allows for innovation and investment in R&D. However, the PPRS has also been criticised for being intransparent in its negotiations and weak in its incentives for cutting prices (Walley, Mrazek and Mossialos, 2005).

In Germany the prices for unpatented, i.e. generic, pharmaceuticals are controlled via a so-called reference pricing scheme that: “[...] aims to contain pharmaceutical expenditure by defining a fixed amount to be paid by the government (or other third-party payer)” (Mossialos and Oliver, 2005). This means that for medicines containing the same substance, similar substances or medicines with comparable efficacy “[...] an upper limit for the costs reimbursable by the sickness funds” (Busse, Schreyoegg and Henke, 2005, p. 345) is imposed. The FJC is responsible for the classification of the substance of these medicines, i.e. grouping together medicines with a similar substance, while the sickness funds set the actual prices.

The price of pharmaceuticals that are new to the German market is dependent on the manufacturer providing evidence that the product in question provides an ‘additional benefit’ in comparison to the medicines currently used in the indication for which the product is licensed. This requirement came into force in January 2011 under the Act on the Reform of the Market for Medicinal Products (AMNOG) (Bundesgesetzblatt, 2010). Before January 2011 pharmaceutical companies in Germany were allowed to set the prices for new products freely. Not only has this freedom been abolished with the introduction of the AMNOG, the extent of ‘additional benefit’ that a new product provides now forms the basis for price negotiations between the manufacturers and sickness funds. While manufacturers can still set the prices for their products, sickness funds no longer have to pay these prices but can negotiate lower prices depending on the extent of the ‘additional benefit’ of a new pharmaceutical. If the FJC finds that the new product does not offer an additional benefit, then the product is assigned to a group under the reference pricing scheme and will only be reimbursed in line with the reference prices set therein.

#### 4.5. ‘The Fourth Hurdle’: Health Technology Assessment (HTA)

In most EU countries acquiring a license for a pharmaceutical product no longer means that it will automatically be covered or reimbursed by a given national health care system. EU member states such as Germany and England have introduced additional control mechanisms in the form of HTA<sup>17</sup> systems in order to decide whether a new medicine is worth its coverage under the national health system (Sorenson, 2009). This means that a market license is no longer a guarantee for pharmaceutical manufacturers to market their products successfully in member states.

Upon acquiring a market license, the pharmaceutical manufacturers embark upon the effort to get their products through different HTA systems across Europe, a process that commentators have labeled the “fourth hurdle” (e.g. Eichler, et al., 2010; Mrazek and Mossialos, 2004) after proving safety, efficacy and quality at the European level. The representations of multi-national pharmaceutical companies in different EU member states work to meet the national requirements for having their products made available in the given health care systems. In order to achieve this they have the same evidence base at their disposal, namely the evidence base that was used for acquiring the marketing authorisation. However, this evidence base needs to be re-worked in order to meet different national requirements. This entails meeting the standards of HTA agencies in different countries.

HTA takes different forms in different national contexts. Sorenson (2009) describes a general trend towards the establishment of HTA agencies to inform decision-making, although this is not a prerequisite for the use of HTA as a decision-making tool. In terms of institutional features, Sorenson (2009) distinguishes between advisory and regulatory bodies. Advisory bodies carry out the assessment of the evidence for the use of a pharmaceutical product whilst having little or no direct decision-making powers on whether the product gets included in the health care benefit basket (Sorenson, 2009). In contrast, regulatory bodies have direct decision-making powers and are involved in the appraisal phase of a product. In addition to their institutional formats, HTAs can also differ in *how* they are carried out, especially with regards to the methods employed to show cost effectiveness, cost-benefit relationships

---

<sup>17</sup> See chapter 1 for a definition of HTA.

or added benefit. These differences will be elaborated on in chapter 5 as they form an important part of HTA policy paradigms in England and Germany.

In the next section I outline the HTA processes and institutions in England and Germany in order to illustrate the procedural and substantive differences between the two systems.

#### **4.6. The HTA Process in England and Germany**

This section provides an overview of the structure, decision-making mandate and procedures of the organisations that are involved in the production of pharmaceutical benefit assessments in England and Germany. These are the Federal Joint Committee (FJC) and the Institute for Quality and Efficiency in Health Care (IQWiG) in Germany and the National Institute for Health and Care Excellence (NICE) in England. While all three institutions have a legal mandate that extends beyond the field of pharmaceutical benefit assessments, for the purpose of answering this thesis' research question I concentrate on explaining how the HTA process for pharmaceuticals works in these institutions. The section focuses on outlining institutional features such as the composition, decision-making remit and procedural workings of the FJC, IQWiG and NICE respectively. While the legal mandate and the possible outcomes of pharmaceutical benefit assessments form part of this outline, the detailed criteria that determine these outcomes are discussed in the next chapter as they represent the ideas that constitute the HTA policy paradigms in Germany and England. Thus, the next section represents an analytical overview of the institutional workings of the FJC, IQWiG and NICE rather than a discussion of how the HTA policy paradigms are applied within these institutions.

Table 4.2. presents a comparative overview of nine of the most important institutional features of the NICE, the FJC and IQWiG. The choice for these institutional features is based on an assessment of what institutional factors might play a role in determining the outcome of pharmaceutical benefit assessments. The most distinctive difference between the FJC/IQWiG and NICE lies in their composition and in their decision-making mandate. While the people who make the final recommendations on a pharmaceutical product at NICE are professionals from the health care arena (NICE, 2009), but not commissioners of services, the individuals who make these decisions in German are members of the previously described bargaining

structures in the German health care system (FJC, 2013). While payers (i.e. CCGs, formerly PCTs) and providers (i.e. professional physician representations) are invited to comment and give evidence on appraisals at NICE, in Germany representations of these entities are the ones who make the decisions. This lack of separation between the bodies that make the decisions and the ones that have to pay for it has led to criticism by other stakeholders in the process, especially by pharmaceutical manufacturers, who question the objectivity of the decision-makers in these processes (Interviewee No. 6, 2013; Interviewee No. 8, 2013). However, because the FJC decision informs the price setting negotiations in Germany rather than the recommendation for the use or non-use of a pharmaceutical product in routine clinical practice, the effect of the involvement of payers and providers in decisions on pharmaceutical benefit assessments might be mitigated by the legal remit of their decisions.

**TABLE 4.2. - Institutional Features of NICE, the FJC and IQWiG: Comparative Overview**

	<b>NICE</b>	<b>FJC</b>	<b>IQWiG</b>
<i>Composition</i>	Appraisal Committee members are appointed for a three-year term and include members from the NHS, patient and carer organisations, academia and representatives of the pharmaceutical industry	Representatives of the self-governing health care system: 5 representatives of the payers (the sickness funds), 5 representatives of the providers (physicians, dentists, hospitals) and 3 impartial members	Salaried employees
<i>Decision-making mandate</i>	Recommendations are binding for NHS organisations (=regulatory body)	Final decision-making power on pharmaceutical benefit assessments (=regulatory body)	Recommendations to FJC, non-binding (=advisory body)
<i>Purpose of assessment &amp; appraisal</i>	To recommend the use or non-use of a medicine in routine NHS practice based on clinical and cost effectiveness criteria	To inform price negotiations between sickness funds and pharmaceutical manufacturers	To assess the pharmaceutical manufacturer's dossier on the 'additional benefit' of a product and make recommendations to the FJC
<i>Who commissions an appraisal?</i>	Commissioned by the Secretary of State for Health, based on topic recommendations by the National Institute for Health Research Horizon Scanning Centre	All pharmaceutical products with a new active substance must be appraised (§135a Social Code Book V)	Commissioned by the FJC
<i>What gets appraised?</i>	Topics referred to by Secretary of State for Health based on criteria such as likely significant benefit to patients, new technology is likely to be at different price	All pharmaceutical products with a new active substance	All pharmaceutical products with a new active substance
<i>Scoping process</i>	Yes	No	No
<i>Patient involvement</i>	Yes. Patient organisations invited to submit comments beginning at the scoping stage.	Yes. No voting rights.	Yes. Patients invited to fill out questionnaire.
<i>Appeals procedure</i>	Yes	No, only via the social courts.	No, only via the social courts.
<i>Publication of appraisal documents</i>	Yes, but not of minutes of decision-making deliberations.	Yes, but not of minutes of decision-making deliberations.	Yes, but not of minutes of decision-making deliberations.

#### **4.7. The HTA Process in England**

NICE was established in 1999 as “[...] a special authority to reduce variation in the availability and quality of NHS treatments and care” (NICE, 2013). This variation in the availability of treatments was a result of the so-called postcode lottery, whereby access to treatment and pharmaceutical depended on where one lived due to the regional health care commissioning authorities, the Primary Care Trusts (PCTs), taking different decisions on how to spend their budgets (Ham, 2009). The establishment of NICE was an outcome of political efforts to reduce the inequitable access to health care that resulted from local variations. Since its inception NICE has undergone several changes in relation to its operational remit. In 2005 its remit was expanded to include guidance on public health and more recently, in 2013, its remit was expanded to the development of guidance for social care practitioners (NICE, 2013).

NICE operates three centres that produce evidence-based guidelines for the NHS. These centres are the Centre for Clinical Practice (responsible for producing clinical guidelines for health care practitioners), the Centre for Public Health (responsible for producing guidelines for public health authorities and practitioners) and the Centre for Health Technology Evaluation (responsible for recommending the use or non-use of pharmaceuticals, medical devices, surgical procedures in routine NHS practice). NICE’s guidance comes in different forms, depending on the centre it is produced in and the topic it pertains to. For the purpose of this thesis, the guidance that comes in the form of so-called technology appraisals (TAs) and that is produced by the Centre for Health Technology Evaluation is the most important guidance.

Based on evidence of clinical and cost effectiveness as well as evidence brought forward by patient and clinical expert groups, TAs recommend the use or non-use of a given pharmaceutical, a diagnostic and surgical procedure or medical devices (NICE, 2014a). TA recommendations are legally binding for the NHS, meaning that local health care commissioners are obliged to fund the treatments that are recommended under the NICE technology appraisal procedure (NICE, 2014a). This legal obligation to fund the services recommended by NICE is different from the guidelines that the Centres for Clinical Practice and Public Health produce because these guidelines are not binding. Regardless of this difference in the binding character of the produced guidelines, NICE issues implementation tools for all of its guidance in order to ensure a wide uptake of its guidance.

Topics for appraisal are referred to NICE by the Department of Health (DH) based on a number of criteria such as the likely impact of a new technology on NHS resources and whether there is a likely inappropriate variation in the clinical practice related to the new or existing technology (NICE, 2008, p. 7)<sup>18</sup>. Before a topic for appraisal is referred to NICE by the DH, the relevant consultees and commentators for the topic are identified and invited to participate in the so-called scoping process (NICE, 2008). Consultees include all groups that might have an interest in or face a direct impact of the new technology under appraisal (NICE, 2008). These groups include patient and carer representations, health professional bodies, manufacturers, the Department of Health, the Welsh Assembly Government and local health boards. Consultees can submit evidence, comment on appraisal documents and appeal the final recommendations made by NICE (NICE, 2008). They can also nominate patient experts or clinical specialists for the consultation process at NICE (NICE, 2008). In contrast to consultees, commentators can only comment on and participate in the appraisal process, but they cannot appeal the final decisions made by NICE. Commentators might include research groups, manufacturers of comparator technologies or organisations such as the MHRA.

The so-called scoping process identifies the disease area, patients, current treatments and clinical practice, likely impact and open questions in relation to the technology for which an appraisal is proposed (NICE, 2008). In this process NICE works closely together with consultees and commentators to get a sense of current clinical practice in a given area. This includes the convention of a scoping workshop in which consultees, commentators and NICE representatives discuss the relevant issues and questions for a given appraisal (NICE, 2008). If the scope suggests that the technology meets the criteria for being appraised, i.e. the impact of a new treatment is likely to be big, and the relevant decision problem and comparator products have been identified, then the DH can refer the topic to NICE for appraisal.

There are two possible routes for technologies to be appraised at NICE, the single technology appraisal (STA) and the multiple technology appraisal (MTA) (NICE, 2012). In a STA a single technology for a single indication is appraised, while a MTA appraises more than one technology for more than one indication. Due to the fact that

---

<sup>18</sup> For the purpose of this thesis the 2008 version of NICE's Guide to the methods of technology appraisal is used. This is because the case studies contained in this thesis were appraised on the basis of this version.

appraisals are frequently conducted for products that are new to the market, meaning that pharmaceutical manufacturers, clinicians and patients might have an interest in getting the product appraised quickly so that their speedy uptake is ensured, the majority of products seem to be appraised under STA procedures. This includes the products that form the basis of the empirical investigation in this thesis, which is why this section explains the STA process in order to illustrate how pharmaceutical benefit assessment works in England<sup>19</sup>.

Once a topic, e.g. a pharmaceutical product, is referred to NICE for an appraisal, the manufacturer of the product submits evidence in relation to the clinical and cost effectiveness of the product in question. At the same time, consultees submit their initial statements and evidence to NICE. In contrast to IQWiG, NICE does not itself evaluate the manufacturer's evidence. Rather it commissions one of several academic research groups, the so-called Evidence Review Group (ERG), to carry out the assessment of the evidence (NICE, 2012). This reflects a common distinction between the assessment and the appraisal stages of HTA processes, assessment being the evaluation of scientific evidence and appraisal being the application of decision-making criteria to this evidence. The ERG carries out the assessment and NICE carries out the appraisal. The ERG assessment and the submissions by consultees and commentators are then collated and referred to a so-called Appraisal Committee (AC) at NICE. Rather than being made up of NICE employees, this committee consists of representatives from the NHS, academia, patient and carer groups and the pharmaceutical industry (NICE, 2009). The committee forms an independent advisory committee to NICE and its individuals sit on the committee for three-year terms. The AC is responsible for producing the recommendations, the appraisal of a product, based on the evidence it receives. During its consultations it also hears evidence from clinical experts and patients who were nominated by consultees.

If the AC recommendations are negative or restrictive they are published in the form of an Appraisal Consultation Document (ACD). Consultees and commentators are then given an opportunity to comment on these recommendations and to submit new evidence. In some cases this additional consultation phase leads to a reversal of the draft recommendations (NICE, 2008). Once the stakeholders have been consulted again, the

---

<sup>19</sup> While the STA and MTA procedures differ in scope and timeline of the appraisal, it is important to note that the methods of appraisal which are touched upon in this section and elaborated on in more detail in the next chapter are the same.

AC makes its final recommendations and publishes them in the form of the so-called Final Appraisal Determination (FAD), which consultees can appeal. If there are no appeals, the FAD forms NICE's guidance on the particular technology at hand. If the AC's recommendations are positive after the initial consideration of the evidence, then the ACD phase of the appraisal is skipped and a FAD is produced instead.

#### **4.7.1. The Dependent Variable: The Outcome of Pharmaceutical Benefit Assessments in England**

There are four possible outcomes of technology appraisals made by NICE: recommended, optimised, only in research and not recommended (NICE, 2014a). If the evidence of clinical effectiveness is convincing and the cost effectiveness ratios are favourable, then an AC usually recommends the product in question for use in the NHS. This is the case when the incremental cost of using a product is no higher than between £20,000-£30,000 per quality-adjusted life year gain (QALYs). QALYs are a measurement of health benefits in relation to life years gained by using a particular medicine or treatment (NICE, 2014b). Together with the concept of the incremental cost effectiveness ratio (ICER) it forms an important part of the HTA paradigm in England and both concepts are discussed in more detail in the next chapter. The ICER refers to the increase in cost that a decision-maker might articulate, "[...] above which a programme would not be acceptable" (Drummond, et al., 2005, p. 43). If an AC decides to recommend a product even though the ICER is higher than £30,000 per QALY gained, it has to make a special case and justify why it still considers the use of the product in question a cost effective use of NHS resources. Similarly, if an AC does not recommend a treatment even though the ICER is below £20,000 per QALY gained, it has to justify why it believes that this is the right course of action.

The 'optimised' category that a pharmaceutical benefit assessment can take refers to instances in which a technology is recommended, but for a smaller sub-set of patients than laid out in the license of a product. This might be the case if there is evidence that a certain group of patients benefits more than others from the use of the medicine in question. The 'only in research' category is applied when the clinical and cost effectiveness evidence is not sufficient yet and the AC therefore recommends the use of the product in research settings, i.e. in clinical trials. Finally, the use of a product is not recommended if there is a lack of sufficient or convincing clinical evidence or its

use would not be considered cost effective, i.e. if the increase in health care cost is not justified in relation to the benefits it provides.

#### **4.8. The HTA Process in Germany**

Since the introduction of the AMNOG in 2011 a so-called early benefit assessment has to be carried out for every pharmaceutical with a new active substance or ingredient. The early benefit assessment is based on a dossier that the pharmaceutical manufacturer has to provide upon introducing a new product to the German market (FJC, 2013a). This dossier includes the manufacturer's assessment of the 'additional benefit' of the product in question and is based on a presentation of the scientific evidence such as RCTs. There is a statutory obligation for the early benefit assessment to present evidence in comparison to the 'appropriate comparator' that is currently used for treatment in the indication for which the new product is licensed (FJC, 2013a). The FJC makes the decision on which alternative medication represents the most appropriate comparator for a given product.

The FJC can commission IQWiG with the production of an early benefit assessment (FJC, 2013a). Without being commissioned, IQWiG cannot carry out an assessment of the benefits of a pharmaceutical product. IQWiG was created in 2004 as an independent scientific body that produces evidence-based reports on a number of topics including pharmaceutical products, surgical procedures, diagnostic procedures, screening programmes, treatment guidelines and disease management programmes (DMPs) (IQWiG, 2013). Its evidence-based reports can take the format of so-called reports, rapid reports, dossier assessments, health information and working papers, depending on what IQWiG is commissioned to carry out by the FJC (IQWiG, 2013). For the purpose of this thesis, this section focuses on the dossier assessments as they are the outcome of the early benefit assessment process.

The early benefit assessment under §35a Social Code Book V (SGB V) takes the form of dossier assessments. Upon receiving the dossier that was submitted to the FJC by the pharmaceutical manufacturer, IQWiG has three months to review and assess the dossier and the evidence contained in it (IQWiG, 2013). Using internal expertise, IQWiG assesses the data contained in the dossiers and can, if necessary, carry out its own research on a given product (IQWiG, 2011a). In order to attain a broader view on a given product's additional benefit, IQWiG involves external clinical experts and patient

groups by inviting them to reply to a questionnaire about the product. At this stage of the early benefit assessment the questionnaire is the only means by which patients and clinical experts are involved. Once the dossier assessment is referred to the FJC for a final decision, an oral and written hearing procedure is carried out by the FJC (FJC, 2013a).

The maximum period for the production of the dossier assessment by IQWiG is three months. After three months the dossier assessment, i.e. the review of the evidence, is referred back to the FJC, which is responsible for making the final decision on a product's additional benefit (IQWiG, 2011a). This decision is prepared in a sub-committee on pharmaceutical products of the FJC before it is referred to the full FJC to appraise the product (FJC, 2013a). The division of labour between IQWiG and the FJC represents the important procedural distinction between the assessment and the appraisal of a medicine's benefits. The assessment reviews the robustness and quality of the scientific evidence whilst the appraisal process includes the application of criteria and value judgements to this evidence in order to reach a conclusion on the benefit of a given product (NICE, 2014b).

#### **4.8.1. The Dependent Variable: The Outcome of Pharmaceutical Benefit Assessments in Germany**

Germany's approach to appraising a pharmaceutical product involves the categorisation of the product's additional benefit. There are six categories of 'additional benefit', which represent the dependent variable of this study. The benefit of a pharmaceutical is defined as "[...] the patient-relevant therapeutic effect, in particular in respect of the improvement in the state of health, the reduction of the duration of the disease, [...] an improvement in the quality of life" (Bundesgesetzblatt, 2010). The issue of 'patient relevance' is at the core of the German HTA policy paradigm and is discussed in more detail in the next chapters. According to Section 5, § 7 of the FJC's Rules of Procedure (G-BA, 2013, pp. 8-9):

[...] the extent of the additional benefit and the therapeutic importance of the additional benefit compared to the appropriate comparator must be quantified as follows, taking into account the severity of the disease:

1. There is a **major additional benefit** if a sustained and large improvement in the therapy-relevant benefit [...] is achieved, which has not previously been

achieved [...], in particular a recovery from the disease, a considerable increase in life, long-term freedom from severe symptoms or extensive avoidance of severe side-effects [...].

2. There is a **significant additional benefit** if a considerable improvement in the therapy-relevant benefit [...] is achieved, which has not previously been achieved [...], in particular a lessening of severe symptoms, a moderate extension in life, an easing of the disease, which is noticeable to patients [...].
3. There is a **marginal additional benefit** if a moderate improvement and not merely a slight improvement in the therapy-relevant benefit [...] is achieved [...], in particular a reduction in non-severe symptoms of the disease or a relevant avoidance of side-effects.
4. There is an **additional benefit, which is not quantifiable** however, because the scientific data base does not permit this.
5. There is **no additional benefit**.
6. **The benefit [...] is less than the benefit** of the appropriate comparator<sup>20</sup>.

The six categories represent the so-called extent of ‘additional benefit’. In addition to the extent of the additional benefit, the appraisal of a product also depends on the ‘probability’ of additional benefit (G-BA, 2013). This ‘probability’ is connected to the quality of the evidence. This means that the extent of additional benefit, i.e. the category assignment a product receives, may be lowered or raised depending on the quality of the presented evidence. The dependent variable, i.e. the outcome of pharmaceutical benefit assessments in Germany, is thus a categorical expression of ‘additional benefit’ that is dependent on the extent of the significance of statistical results on the one hand and the quality of the presented evidence on the other. How this categorisation is operationalised in practice forms a discussion that is presented in the empirical chapters of this thesis. For now, suffice it to say, that the six categories of additional benefit serve the purpose of informing price negotiations between the sickness funds and pharmaceutical manufacturers. In theory, the higher the category that an early benefit assessment results in, the higher the price that the pharmaceutical manufacturers can demand from the sickness funds. Vice versa, if a product receives a low category of additional benefit, the sickness funds will not be as willing to pay a high price for the product. In the German context the concept of thresholds is employed to describe thresholds between different categories of additional benefit, which impact on the price negotiations between the sickness funds and the pharmaceutical manufacturers.

---

<sup>20</sup> Highlights added by the author of this thesis.

#### **4.9. Institutional Features of NICE, the FJC and IQWiG: Comparative Summary**

The preceding sections provided an overview of the pharmaceutical benefit assessment processes in England and Germany. The institutions that carry out the benefit assessment both produce evidence-based guidance. As such their mandates are to inform health care decision-making based on the available scientific evidence in a given field. However, the way they carry out this mandate differs procedurally and substantively. It differs procedurally in the commissioning of the evidence review to outside groups and in the procedural steps that characterise the HTA process. Substantively, the HTA processes differ in the way the benefit assessment is operationalised, especially with regards to the way the concept of thresholds is employed.

Procedurally, the main difference between the HTA processes lies in the fact that two bodies (FJC and IQWiG) are involved in the assessment and appraisal process in Germany while there is only one in England (NICE). However, the significance of this should not be overestimated as it is connected to the distinction between the assessment and the appraisal stages of a HTA process rather than to a substantially different idea on how to carry out HTAs. That is to say, Germany chose to create an independent organisation, IQWiG, to carry out the evidence assessment while policy-makers in England did not do so. Yet, in England the assessment part of an HTA process gets ‘outsourced’, i.e. NICE commissions an ERG with this task. This is comparable to the FJC commissioning IQWiG to carry out the dossier assessment in Germany. Thus, the role played by IQWiG in Germany is akin to the role played by the ERG in England. In both cases, the tasks of assessment and appraisal are separate and the final decision-making occurs outside of the bodies charged with the evidence assessment.

Another procedural difference lies in the number of steps that make up the HTA process. At NICE, the number of steps is large and includes the scoping process, the evidence review, consultations in both of these phases, the ACD, further consultations, the FAD and finally the appeals procedure. What is important about this observation is that consultees and commentators are involved not just at one stage of the appraisal process, but at each stage. They have an opportunity to comment in the scoping phase, they submit evidence during the consultation phases, they can comment and submit new

evidence at the ACD phase and they can appeal the final decision. This means that the opportunities for involvement at NICE are vast.

The number of steps in the German HTA process is more limited by comparison. There is no formalised scoping process, but only an opportunity for the pharmaceutical manufacturer to apply for an advisory meeting with the FJC ahead of an early benefit assessment. There are no opportunities for stakeholders to comment on FJC draft proposals. The FJC hearing is based on the dossier assessment produced by IQWiG, which provides a categorisation of the benefit, but does not give an indication of whether the FJC is likely to follow this categorisation in its appraisal. As such there is a procedural difference in the number of steps that stakeholders are involved in and in the fact that hearings are carried out on the *evidence review* rather than on the initial *appraisal decision* pertaining to a product.

Substantively, the HTA processes in Germany and England differ in the purpose for which they are conducted and in the way ‘benefit’ is operationalised. In Germany the early benefit assessment is carried out in order to inform price negotiations between the sickness funds and the pharmaceutical manufacturers. Whilst the FJC has a decision-making remit that extends to the inclusion or exclusion of a service in the health care benefit basket, the default position for new pharmaceuticals is that they will be reimbursed, the question is just at which price. This is a substantively different paradigmatic situation from that in England where the default position for new pharmaceuticals is that they will only be covered if they are clinically and cost effective.

The substantively different purposes that HTAs serve might explain the differences in how ‘benefit’ is being operationalised in Germany and England. In Germany pharmaceutical benefits are categorised into one of six categories in order to support negotiating parties in compromising on an acceptable price. Thresholds are used when it comes to distinguishing between these categories, although chapter 7 and 8 show that these thresholds are neither fully transparent nor accepted amongst stakeholders. In England pharmaceutical benefits are viewed in relation to their costs. This suggests that Germany’s HTA paradigm operates on the presumption that the value or benefit of a pharmaceutical, even if it is low, can be reflected in the price the payers pay, while the English HTA paradigm operates on the idea that a threshold exists at which the value or benefit ceases to be sufficiently beneficial in relation to its costs and thus its uptake is no longer justified. To put it differently, the German HTA paradigm is characterised by the idea that there are thresholds for the ‘right’ price of a

pharmaceutical, while the English HTA paradigm is characterised by the idea that there are thresholds for the inclusion or exclusion of a pharmaceutical product from use in the NHS.

#### **4.10. Conclusion**

This chapter introduced the reader to the institutional features that characterise the HTA paradigms in Germany and England. It did so by summarising the most important procedural and substantive differences between the HTA processes in Germany and England. On a procedural level these include the different bodies to whom the assessment task of an HTA is ‘outsourced’ and the number of formalised procedural steps that need to be followed in order to protect HTA outcomes from appeals. On a substantive level the differences include a different understanding on what the purpose of HTAs is and how they are operationalised.

Whilst both the FJC and NICE employ thresholds to aid decision-making, the purpose of these thresholds and how they are operationalised is very distinct. In Germany thresholds are used to assign benefit categories and inform price setting negotiations whilst in England they are used to inform the inclusion or exclusion of a pharmaceutical product for use in NHS practice. Whether or not this distinction matters in determining the final outcome of HTA processes is explored in the analytical chapters of this thesis. However, before exploring this question, I further elaborate on the policy paradigms of pharmaceutical benefit assessments with regards to their ideational basis in the next chapter.

## Chapter 5

### The Policy Paradigms of Pharmaceutical Benefit Assessments in England and Germany

#### 5.0. Introduction

This chapter offers an analytical overview of the pharmaceutical benefit assessment policy paradigms in Germany and in England. The policy paradigms are discussed by reference to a) the legislative framework that provides the construct for the health care systems and b) the methods guidelines of the National Institute for Health and Care Excellence (NICE), the Federal Joint Committee (FJC) and the Institute for Quality and Efficiency in Health Care (IQWiG). As such this chapter, along with the previous chapter, identifies the broad HTA paradigms in England and Germany. However, the identification of the paradigms in the relevant statutory materials does not offer insights into how the paradigms operate, i.e. what their role is in ‘normal’ HTA decision-making. As Cartwright and Hardie (2012) and Majone (1989) highlight, in policy-making evidence requires careful judgement and interpretation. While the identification of the paradigms in this chapter provides an initial impression of how evidence questions might be approached in England and Germany, this is supplemented by the empirical analysis that follows in the next chapters in order to understand how the paradigms are operationalised in practice.

For ease of understanding this chapter is divided into parts that reflect how policy paradigms are operationalised for the purpose of this study. In comparison to the previous chapter the principles discussed in these sections are more ideational in kind. In sections 1-4 of this chapter I lay out the values on which the health care systems in England and in Germany are built. I then discuss the political and institutional purpose of benefit assessment which is different in Germany and in England. In sections 6-8 I discuss the principles and values that guide HTA decision-making as they are specified in the methods guidelines before outlining the criteria of HTA decision-making. In section 9 I analyse the case of Apixaban, a pharmaceutical product for the prevention of thromboembolic events after hip- or knee replacements, as a case that represents a comparatively straightforward application of the HTA paradigm in both Germany and England. Finally, I offer some concluding remarks that provide guidance for the interpretation of the empirical findings in the next chapters.

Table 5.1.<sup>21</sup> provides an overview of the values and principles contained in the statutory framework and methodological guidelines, which shape the work of NICE, the FJC and IQWiG. Together, these values and principles constitute the broad HTA paradigms in England and Germany. The identification of the HTA paradigms resulted in a number of findings. Firstly, the health care paradigms give rise to tensions that need to be resolved when HTA paradigms are established in their routine, i.e. normal, application. For example, both the English and the German paradigms contain principles that relate to the breadth of health services that should be provided whilst ensuring an efficient, or cost effective, use of resources without making it clear how decision-makers should balance these principles. This balancing act occurs during normal decision-making in which different rules and criteria are applied. The differences between the German and the English HTA paradigm are especially notable when it comes to the use of thresholds even though the underlying idea of HTAs appears to be similar, that is costs or prices of pharmaceuticals need to be justifiable in relation to their clinical benefits. In England this idea is operationalised by using incremental cost effectiveness ratios (ICERs), whereas in Germany it is operationalised by reference to ‘patient relevance’ and benefit categories.

---

<sup>21</sup> The table distinguishes between *principles* and *criteria* of HTA decision-making. The principles of decision-making (number three) are the values that are embedded in HTA decision-making, i.e. that arise from the wider paradigm, while the criteria represent the way in which these principles are supposed to be translated into practice. The criteria for decision-making (number four) are the means by the values are safeguarded, they are the rules of the game under which the paradigm operates in normal decision-making. For example, in the case of NICE equality and non-discrimination represent the values that guide NICE’s decision-making whilst the application of social value judgements and end-of-life-criteria are the means by which they are translated into practice.

**TABLE 5.1. - The HTA Paradigms in England and Germany**

<b>Features of the HTA paradigm</b>	<b>NICE (Main HTA decision-making body)</b>	<b>FJC (Main health care decision-making body)</b>	<b>IQWiG (Body that carries out HTAs)</b>
<i>1) Principles &amp; values within the wider health care context (= broad policy paradigm)</i>	<u>NHS Constitution (Department of Health, 2013a):</u> - <b>Comprehensive</b> service, ensuring <b>equality &amp; non-discrimination</b> - <b>Clinical need</b> - Service to ensure <b>excellence &amp; professionalism</b> - <b>Value for taxpayers' money</b> - <b>High quality care</b>	<u>German Constitution (BMJV, 1949):</u> - <b>Social/welfare state principle</b> (§20 & 28) <u>Social Code Book V (BMJV, 2013):</u> - <b>Solidarity</b> (§ 1) and <b>individual responsibility</b> - Provision of health care service based on <b>'generally accepted state of medical knowledge'</b> (§ 2) - <b>Efficiency</b> (§ 12) - In cases of severe, rare and/or life-threatening illnesses: <b>Right to health care intervention</b> even when 'generally accepted state of medical knowledge', i.e. evidence, is not available or not conclusive ('Nikolaus' court ruling) (Das Bundesverfassungsgericht, 2005)	<u>German Constitution (BMJV, 1949):</u> - <b>Social/welfare state principle</b> (§ 20 & 28) <u>Social Code Book V (BMJV, 2013):</u> - <b>Solidarity</b> (§ 1) and <b>individual responsibility</b> - Provision of health care service based on <b>'generally accepted state of medical knowledge'</b> (§ 2) - <b>Efficiency</b> (§ 12) - In cases of severe, rare and/or life-threatening illnesses: <b>Right to health care intervention</b> even when 'generally accepted state of medical knowledge', i.e. evidence, is not available or not conclusive ('Nikolaus' court ruling) (Das Bundesverfassungsgericht, 2005)
<i>2) Purpose of HTA process as set out in methodological guidelines (=narrow policy paradigm)</i>	- To <b>appraise health benefits and costs</b> of medical interventions (NICE, 2008)	- To make <b>transparent and legally sound decisions</b> that reflect the <b>'generally accepted state of medical knowledge'</b> (BMJV, 2013)	- <b>To support FJC in making its decision</b> about the coverage of pharmaceuticals under SHI schemes <b>by assessing the therapeutic benefit of a new medicine</b> (IQWiG, 2013)

<p>3) <i>Principles &amp; values that guide HTA decision-making (= narrow policy paradigm)</i></p>	<ul style="list-style-type: none"> <li>- <b>Consistency</b></li> <li>- <b>Equality</b></li> <li>- <b>Non-discrimination</b></li> </ul> <p>(NICE, 2008)</p>	<ul style="list-style-type: none"> <li>- <b>Transparency and legal soundness</b></li> <li>- Decisions should reflect ‘<b>generally accepted state of medical knowledge</b>’</li> <li>- Consider issues of <b>quality and health care provision</b></li> <li>- Consider <b>interests of stakeholders</b></li> <li>- <b>Efficiency</b> (BMJV, 2011)</li> </ul>	<ul style="list-style-type: none"> <li>- <b>Independence</b></li> <li>- <b>Patient-oriented</b></li> <li>- <b>Transparency</b></li> <li>- <b>Evidence-based medicine</b></li> </ul> <p>(IQWiG, 2011a)</p>
<p>4) <i>Criteria and methods of HTA decision-making (Instrument settings of the paradigm)</i></p>	<ul style="list-style-type: none"> <li>- <b>Clinical &amp; cost effectiveness</b></li> <li>- <b>End-of-lifetime considerations</b></li> <li>- <b>Social value judgements</b></li> <li>- <b>Innovation</b></li> <li>- <b>Assessment across disease areas</b> (NICE, 2008)</li> </ul>	<ul style="list-style-type: none"> <li>- <b>Therapeutic benefit (=patient relevant therapeutic effect)</b></li> <li>- <b>Medical necessity</b></li> <li>- <b>Efficiency</b></li> <li>- <b>Assessment within disease areas</b></li> </ul> <p>(G-BA, 2013)</p>	<ul style="list-style-type: none"> <li>- <b>Evidence-based medicine</b></li> <li>- <b>Therapeutic benefit (= patient relevant therapeutic effect)</b></li> <li>- <b>Assessment within disease areas</b></li> </ul> <p>(IQWiG, 2011a)</p>

## **5.1. Principles and Values within the wider Health Care Context**

According to the World Health Organization (WHO) the purpose of a health care system is to promote, maintain or restore health (WHO, 2014). This purpose lies at the core of health care systems around the world. However, the means by which countries seek to fulfill the purpose are varied. The definition of health, the financing of health care and the expectations of populations and policy-makers are but a few of the factors that depend on national values and principles. In order to understand pharmaceutical benefit assessment outcomes it is therefore necessary to understand the wider values and principles that guide health care in a given country.

## **5.2. England: Principles and Values in Health Care**

The values and principles that form the core of the National Health Service (NHS) in England can be found in the so-called NHS Constitution (Department of Health, 2013a). This document distinguishes between the principles and values of what the NHS does. There are seven principles for the services that the NHS provides:

1. NHS services are comprehensive and available to all;
2. Access to NHS services is based on clinical need and not ability to pay;
3. Excellence and professionalism is aspired to when providing services;
4. Patients are at the heart of NHS services;
5. The NHS works across service boundaries and with other services in the interest of its patients;
6. The NHS seeks to provide the best value for taxpayers' money and the most effective and fair use of finite resources;
7. The NHS is accountable to the public and patients (Department of Health 2013a).

Embedded in the above principles is a myriad of social values. Principles one and two reflect the values of equality and non-discrimination. The NHS Constitution specifies principle one by stating that NHS services are:

“[...] available to all irrespective of gender, race, disability, age, gender orientation, religion, belief [...]. It [the service] has a wider social duty to promote equality through the services it provides” (Department of Health, 2013a, p. 3).

In principle one *comprehensiveness* is highlighted as a principle in its own right, but the rest of the principle focuses on making NHS services available to all. Given the NHS' long history of making difficult prioritisation decisions within the context of limited health care budgets (Robert, 2003), the question arises whether *comprehensiveness* is defined in terms of a *comprehensive access* to services rather than access to *comprehensive*, i.e. all-encompassing, *services*.

From a comparative point of view this is an important question because it is an example of how the principle of comprehensiveness might be laid out differently in different HTA policy paradigms. *Comprehensive access* is different from *comprehensive services*. The latter suggests that health care services are likely to include most of what is medically and technologically possible, but it does not mandate who will have access to these services. *Comprehensive access* on the other hand guarantees that all members of society have access to the health care services provided, but it does not mandate that these services have to be comprehensive. Whilst a detailed discussion of this issue is beyond the scope of this thesis, the distinction between comprehensive access and comprehensive services may help explain why rationing and prioritisation measures are deemed acceptable in the English context despite the fact that comprehensiveness is manifested in principle one of the NHS Constitution. It might also help explain why the centrality of cost effectiveness thresholds in the English HTA paradigm, which is further discussed in a later section of this chapter, is not perceived to be at odds with the wider health care paradigm.

In addition to the social values of equality and non-discrimination contained in principles one and two, principles five-seven indicate a high regard for the values of professionalism, transparency and accountability in the English context. Based on the distinction between procedural and substantive values made by Kenny and Joffres (2008), the latter two can be labelled procedural values in that they connote a responsibility by the NHS towards the public, the patients and the taxpayers. This responsibility needs to be fulfilled by ensuring transparent, professional and accountable decision-making procedures, as stated implicitly in principle five-seven. As outlined in the previous chapter, the technology appraisal process at NICE reflects these principles of the English health care paradigm by anchoring transparency and accountability as important features in the decision-making process. Consultees are engaged at every step of the appraisal process, all relevant documents are published and there is a clear appeals process in place, which ensures accountability.

It is interesting that the NHS Constitution includes a principle (principle six) that explicitly spells out the fact that the NHS is faced with finite resources. Principle six connects the reference to finite resources with the values of fairness and the need to make decisions that represent best value for taxpayers' money (Department of Health, 2013a). Despite the principle being comparatively short in length, it gives insight into the challenges faced by the NHS and the tensions that the principles of the NHS Constitutions give rise to. That is to say that principle six might not always sit easily with the other principles that the NHS strives to promote and maintain. However, by including principle six, policy-makers and NHS practitioners have made it explicit that the NHS resources are finite and that tough decisions, which balance the interests of patients and the wider public have to be made. As highlighted in section 5.3., a comparative acknowledgement of the finite nature of resources is lacking in Germany.

The fact that the challenge of finite resources is explicitly mentioned in a guiding document provides a paradigmatic framework that decision-makers can fall back on. In this sense *explicitness* itself, i.e. being explicit about challenges, can be interpreted as a value that is contained in the English health care system. By being explicit about the fact that the NHS is facing finite resources, the principles of the NHS Constitution can be interpreted as a paradigmatic framework that seeks to mitigate against high expectations that patients might have and offer decision-makers a value construct to guide their decisions.

The NHS Constitution (Department of Health, 2013a) goes on to specify a number of additional values that are important in the NHS context. These values include compassion, respect, dignity, improvement of life, quality of care and non-discrimination as values that should be at the heart of everything the NHS does. However, for the purpose of this thesis the focus lies on the principles and values discussed in the previous paragraphs as they reflect procedural and substantive values that play an important role in NICE's technology appraisals. That is not to say that values such as compassion, respect and dignity do not play an important role. However, these values are more important in the context of the every-day clinical services provided by the NHS and the relationship between NHS staff and patients.

In summary, the broad paradigm of the health care system in England is characterised by a number of social values that are not limited to the health care context. Equality and non-discrimination are values that go beyond the health care context, but they may be especially important to promote in health care as individuals are more

vulnerable in this setting. **Equality, non-discrimination** and **comprehensive access** to health care services feature prominently in the NHS' guiding framework (Department of Health, 2013a). These, along with the requirement for **fairness and value for money** in making funding decisions, can be viewed as the substantive ideas that provide a construct for the English health care system. On the procedural side, the NHS Constitution stipulates **professionalism, transparency and accountability** (Department of Health, 2013a). Again, these are values that are not limited to the health care context, but guide other areas of public policy too.

The principles and values embedded in the NHS Constitution are a reflection of the institutional structure of the English health care system in that ideas such as value for money, transparency and accountability are important in systems where health care services are paid for by a collective societal entity, in this case the taxpayer community. Most importantly, the acknowledgement that health care resources are finite is an important part of the health care paradigm that is reflected in the 'rules of the game' in the HTA process. Such an acknowledgement does not feature in the German health care paradigm.

### **5.3. Germany: Principles and Values in Health Care**

The principles and values that the German health care system is built on can be found in the Constitution of the Federal Republic of Germany (BMJV, 1949) as well as the legislative framework that mandates health care, the Social Code Book V (BMJV, 2013).

Germany's Constitution (BMJV, 1949) is important for understanding the underpinnings of the health care system in that it contains the social/welfare state principle (Art. 20 & 28), which stipulates that the German state is a welfare state. The welfare state principle forms the basis of welfare and health care provisions. The operationalisation of principles and values in respective areas can be, and has been, challenged in front of the courts in Germany. There have been important Federal Constitutional Court rulings such as the so-called Nikolaus-decision<sup>22</sup> in 2005 in which the state's obligations have been interpreted generously in favour of patients who are seeking access to health care services (Das Bundesverfassungsgericht, 2005). As we

---

<sup>22</sup> Named after the day of Saint Nicholas which is traditionally celebrated in Germany on 6 December and which was the date of the court ruling in question.

shall see, in the Nikloaus-decision the Federal Constitutional Court decided that patients with rare or life-threatening conditions are entitled to specific treatments even if the clinical benefits of that treatment are not conclusive. The constitutional welfare state principle has wide-ranging implications for the way in which Germany's health care system functions. On the institutional level this includes a large role that the courts play in defining the state's obligations, and the limits thereof, under the welfare state principle.

Moving from the constitutional principle of the welfare state to the principles found in the statutory framework, the values of solidarity, efficiency, individual responsibility and medical/scientific knowledge are laid out as guiding principles within the health care context (BMJV, 2013). In Article 1 of the Social Code Book V the statutory sickness insurance is defined as a solidaristic community that is charged with the task of maintaining, restoring and improving the health of the insured members. Article 3 of the Social Code Book V extends the principle of solidarity to the way in which the health care system is financed, i.e. through shared contributions by employees and employers and through risk-pooling between and across insurance funds.

Individual responsibility is anchored in Article 1 of the Social Code Book V (BMJV, 2013). According to this article individuals have a responsibility to live healthily and take preventive measures in order to avoid ill health whilst sickness funds have a responsibility to promote healthy living through the services they provide (BMJV, 2013). Buyx (2008) highlights that personal responsibility is frequently mentioned as a possible criterion for denying patients access to a specific health care service in the academic and political debates about rationing and prioritisation, for example when a disease is thought to be brought about by a particular lifestyle choice such as smoking or excessive food consumption. However, due to ethical concerns associated with the difficulty of demonstrating causality between lifestyle choice and a disease, health care and policy decision-makers shy away from pursuing personal responsibility as a criterion for prioritisation decisions (Buyx, 2008).

It is noteworthy that Germany's statutory framework contains individual responsibility, not just as a footnote or a by-line, but as a principle in a prominent position. However, personal responsibility is not used as a criterion to decide which patients have access to services or what services are included in the health care benefit basket. Personal responsibility appears to be better explained in terms of the importance of the principle of solidarity in which members of the solidaristic community

understand that their right to services comes with a commitment to individual responsibility. The fact that personal responsibility is neither referred to in the legislative framework setting out HTA decision-making criteria nor considered in any of the empirical cases analysed as part of this thesis supports the understanding of personal responsibility as being connected to the principle of solidarity rather than a stand-alone value.

The Social Code Book V (BMJV, 2013) sets out efficiency and the ‘generally accepted state of medical knowledge’ as principles to guide health service provision and decision-making in Germany. According to Article 12 services have to be sufficient, appropriate and efficient (BMJV, 2013). However, the article does not specify what is considered efficient, which suggests that the operationalisation of the term is at the discretion of the self-governing health care bodies in Germany. Moreover, the ‘generally accepted state of medical knowledge’ is not specified. As section 8 of this chapter shows, in the context of pharmaceutical benefit assessments this principle is operationalised by using methods associated with the field of evidence-based medicine.

Finally, for cases of severe, rare and life-threatening diseases, Article 2 of the Social Code Book V (BMJV, 2013) contains a right for patients to access a treatment even if it is not part of what is usually funded by the sickness funds or if the efficacy evidence for the treatment is not conclusive. This article is a result of the previously mentioned Nikolaus-decision of 6 December 2005 in which the Federal Constitutional Court overruled previous social court rulings with regards to a patient with a rare disease called Duchenne muscular dystrophy (Das Bundesverfassungsgericht, 2005). There is currently no generally accepted and evidence-based treatment for this disease. The patient made an application for his sickness fund to cover the cost for a so-called bioresonance therapy which he had already been receiving but which his parents had previously paid for (Das Bundesverfassungsgericht, 2005). The sickness fund refused this application on the grounds that there was no causal evidence to suggest that bioresonance therapy would improve the patient’s health status. After the social courts upheld the sickness fund’s decision, the Federal Constitutional Court overturned the social courts’ ruling and ruled that the sickness fund’s decision was a breach of the patients’ basic rights under Article 2 (right to free development of personality and right to life and physical integrity) of the German Constitution (BMJV, 1949). The Federal Constitutional Court also connected this breach of rights to the state’s obligation under the welfare state principle.

Article 2, paragraph 2 of the Social Code Book V (BMJV, 2013) reflects the ruling in the Nikolaus case by granting patients the right to treatment even if there is no evidence of efficacy of a treatment. Even though the article pertains to life-threatening and rare diseases for which no alternative treatments are available, the Nikolaus-decision has significant implications for the ability of health care decision-makers to restrict access to treatments on the grounds of efficacy and efficiency. Arguably, the implementation of rationing and prioritisation policies has been made more difficult as a result of the Nikolaus-decision, which is now firmly anchored in the German health care paradigm.

In summary, the values that frame health care in Germany include a constitutionally embedded **welfare state principle** (BMJV, 1949), the principles of **solidarity and personal responsibility** (BMJV, 2013), the values of **efficiency and medical knowledge** (BMJV, 2013) in the form of evidence-based medicine as well as a **right to health care treatment** that overrides evidence-based medicine in severe cases. The ideational base of the Social Code Book V is connected to a number of values frequently associated with the welfare and health care areas rather than with area of public life at large. Altenstetter and Busse assert that the normative ideas underlying statutory health insurance policy-making in Germany “[...] may include competitive elements but are mediated by the influence of the three Ss – solidarity, self-governance, and subsidiarity (i.e., leaving decisions to the smallest capable unit)” (Altenstetter and Busse, 2005, p. 138).

#### **5.4. Comparative Views**

The preceding sections identify the main features of the English and the German health care paradigm. The identification thereof is important because it gives rise to the HTA paradigms that are identified in the next sections. What is striking from a comparative point of view is that in the German health care context one finds little to no reference to wider social values such as equality and non-discrimination. In contrast, these social values play a big role in the NHS Constitution and NHS decision-making generally. A possible explanation for this might be the German Constitution, which guarantees rights in all areas of public life. The Constitution includes the right to equality, non-discrimination and freedom of expression (BMJV, 1949). By

constitutionally guaranteeing rights such as equality and non-discrimination, these values do not have to be reiterated in every piece of legislative guidance.

Another possible explanation for the absence of any reference to wider social values in the German paradigm is connected to the lack of acknowledgement that health care resources are finite. Acknowledging that resources are finite implies an acknowledgement that not everything can be done for everyone and that difficult choices will have to be made. In an effort to make these choices fair, legitimate and acceptable, principles such as equality and non-discrimination become more important. This might explain why the NHS Constitution emphasises the importance of non-discrimination and equality. To put it differently, if it is not acknowledged that resources are finite, then it is less important that they are distributed equally because there are enough resources for everyone to benefit.

There are flaws to the above attempts at explaining the absence of wider social values from the German health care paradigms, the most obvious one being the implication that the need for equality and non-discrimination is greater when resources are finite. This implication is flawed because even in a hypothetical system of infinite health care resources decisions would have to be made about how to distribute the resources in order to ensure that everyone, and not just certain members of society, benefit from them. A further elaboration on these points is beyond the scope of this thesis. The main point for the discussion that follows is that the English health care paradigm specifically acknowledges the finite character of health care resources while the German one does not. At the same time wider social values that are not just limited to the health care arena can be identified in the English paradigm, but not in the German one. Whether or not these features are associated with one another – i.e. whether an acknowledgement of finite resources necessitates the adherence to wider social values – is a question that could be addressed in future research.

The overview of values embedded in the health care paradigms also shows that procedural values such as transparency and accountability are not mentioned in the German context. The principles referred to in the German context are more substantive in kind, whereas in England they are both procedural and substantive. The lack of procedural values in the German context may be explicable with reference to the self-governing health care system which implies that procedural issues need to be dealt with at self-governing level rather than at state level.

It is an interesting empirical observation that the lack of procedural values at paradigm level is congruent with a frequently voiced criticism by stakeholders about the lack of transparency in health care decision-making in Germany (e.g. Interviewee No. 6, 2013; Interviewee No. 8, 2013; Interviewee No. 9, 2013). This observation is discussed in more detail in chapters 6, 7 and 8. For now, suffice it to say that there appears to be a difference in the importance attributed to procedural values in the German and English paradigms. It suggests that in comparing the German and English systems of pharmaceutical benefit assessments one faces two cases that place a different emphasis on substantive and procedural values. Whether or not this difference is reflected in pharmaceutical benefit assessment processes and whether it impacts on the outcome of assessments will have to be addressed when comparing the empirical results with the discussion of paradigms presented in this chapter.

A common characteristic of the English and the German HTA paradigms is that they give rise to tensions that have to be resolved at the decision-making level, i.e. during HTA decision-making. At their core, HTA processes are about the identification of clinically effective vs. non-effective treatments and/or about the identification of cost effective vs. cost ineffective treatments. In the English health care paradigm, the finite nature of resources is acknowledged whilst, at the same time, the principle of comprehensive services is firmly anchored in the NHS Constitution. Similarly, the German health care paradigm includes the principle of efficiency, whilst the Nikolaus-ruling resulted in a statutory paragraph that potentially challenges attempts to adhere to the principle of efficiency.

Neither the English nor the German paradigms provide decision-makers with guidance on how to resolve the tensions that arise, i.e. how to provide comprehensive services when resources are finite or how to provide efficient services when severely ill patients have a right to treatment for which there is no evidence that it is effective, let alone efficient. These tensions need to be resolved at the 'normal' decision-making level. They support the theoretical argument made in chapter 2 that it is not enough to identify the broad features of a paradigm when trying to understand their effect on empirical phenomena. Paradigms acquire meaning and articulation when they are operationalised in practice and the empirical chapters that follow provide initial insights into how the above tensions, among others, are resolved in practice. While paradigms set boundaries for what is considered possible in a policy area, the outlined tensions suggest that there is room for how these boundaries are established in practice.

Before turning to the purposes and goals as institutional features of HTA paradigms, I briefly raise the question of whether the above tensions result in a situation in which certain values are given more weight than others. This question emerges from the Nikolaus-case (Das Bundesverfassungsgericht, 2005) in Germany, which appears to suggest that the right to treatment trumps the principle of evidence-based medicine anchored in the Social Code Book V and the guidelines of the Federal Joint Committee (FJC). If this is the case, then the decision problem becomes one of how decision-makers adhere to guiding values such as efficiency without risking patients invoking their rights under Article 2, paragraph 2 of the Social Code Book V (BMJV, 2013). In the context of pharmaceutical benefit assessments, the empirical analysis undertaken in this thesis suggests that a core and a periphery of values exist. The core is the cost effectiveness thresholds in England and patient relevance in Germany, both of which appear to be weighted more than other principles contained in the decision-making paradigms. The empirical evidence thus suggests that certain concepts trump others, a finding that adds another tier of complexity to the analysis of policy paradigms.

### **5.5. Purpose and Goals of HTA Processes**

The goals HTA processes are designed to achieve are important features of HTA policy paradigms as they inform the way the processes are structured and provide explanations for the methods that are employed in these processes. The purposes that HTA processes serve in different contexts can thus be viewed as an integral part of the policy paradigms that guide them.

The previous chapter outlined that HTA processes in England inform decisions on the inclusion or exclusion of pharmaceutical products in routine NHS practice whereas in Germany they serve as the basis for price negotiations between the pharmaceutical manufacturers and the sickness funds. This institutional difference represents different ideas about what purposes HTA processes can and should serve. This in turn explains the difference in how the purpose of HTA processes is conceptualised in the methodological frameworks for the FJC, IQWiG and NICE. In the next section I turn to this conceptualisation in order to show how different understandings about a policy idea, that of HTA, are reflected in the way the goals of HTA processes are framed in institutional guidelines.

### 5.5.1. England and Germany: Purpose of HTA Processes

The guiding document pertaining to pharmaceutical benefit decisions at NICE is the ‘Guide to the methods of technology appraisal’ (NICE, 2008). For the purpose of this thesis, the 2008 version of this guide is used as this version informed the appraisals that are analysed as part of this research. According to this guide, the purpose of NICE appraisals is “[...] to appraise the health benefits and the costs” (NICE, 2008, p. 7) of technologies, e.g. of pharmaceutical products. As will be shown in the next sections, NICE operationalises health benefits as clinical effectiveness and costs as cost effectiveness in relation to the opportunity costs of services that are possibly being displaced by a positive recommendation from NICE (NICE, 2008). This is important to note because it shows that health benefits and costs are viewed in relation to one another within the HTA paradigm rather than viewing the two as separate stand-alone ideas.

The FJC identifies its purpose as making transparent and legally sound decisions that reflect the generally accepted state of medical knowledge (G-BA, 2013a). With regards to pharmaceutical benefit assessments it does so by carrying out early benefit appraisals of a new medicine in order to inform pharmaceutical price setting in Germany. The inclusion of the ‘generally accepted state of medical knowledge’ as a criterion to be considered when carrying out its tasks is one that reoccurs in a number of the guiding documents of HTA processes and it reflects the principles in the Social Code Book V. The reference to ‘making legally sound decisions’ can be attributed to the big role that the courts in Germany play with regards to the legality of health care decision-making. That is to say that the German system is characterised by a comparative lack of institutional appeals structures (Landwehr, 2009) and the most common way to appeal FJC decisions or an individual funding decision by a sickness fund is to take legal action in court. Thus, in making its decisions, the FJC is particularly aware of making decisions that are legally sound. As one representative phrased it in an interview: “[...] in Germany [...] in contrast to other countries, everything can be taken to court and this sword of Damocles of judicial appeals hangs over all of our decisions” (Interviewee No. 16, 2013, pp. 5-6).

The fact that the procedural value of transparency is mentioned by the FJC suggests that this is a means by which the Committee seeks to ensure the aforementioned legal soundness of their decisions. How this value is perceived and operationalised is the subject of a later part of this thesis. Suffice it to say at this point

that the reference to transparency in a section pertaining to the purpose of the FJC's work is noteworthy in that it indicates that procedural values play a role in the FJC's work despite their comparative under-representation in the value framework discussed in the previous section. Finally, IQWiG defines its purpose as supporting the FJC in its decision-making by assessing the therapeutic benefit of a new medicine (IQWiG, 2011a). This reflects IQWiG's institutional role as the assessment rather than the appraisal body in Germany.

While the purpose of HTA processes in Germany includes a reference to transparency, a procedural value, a similar reference is lacking in the methodological guidelines of NICE. However, the lack thereof in the English context should not be overstated as it seems to be a result of semantic and structural lay-out decisions rather than of decisions that are paradigmatically significant. As the previous section on the wider values in health care in England highlighted, procedural values play a large role in health care decision-making in general. The next sections show that this is reflected in the principles and criteria of HTA decision-making.

In summary, the purpose and task of both NICE and the FJC is the *appraisal* of pharmaceutical products. However, the difference between the organisations pertains to what aspects in relation to the scientific evidence on a product are being appraised and to what end. In England, health benefits and costs are appraised with the purpose of recommending the use or non-use of a pharmaceutical product in routine clinical practice. By contrast, in Germany the question of costs is left aside. Only the 'therapeutic benefits' of a new medicine are appraised with the purpose of informing pharmaceutical price setting in Germany. The aspects that are considered and the ends of an appraisal can be illustrated schematically in the following way:

**England:**

Assessment & appraisal of health benefit + assessment & appraisal of costs →  
Recommendation of the use or non-use in routine clinical practice

**Germany:**

Assessment & appraisal of therapeutic benefits → Pharmaceutical price setting  
to reflect the added benefit of a product

The significance of the similarity of *appraising* benefits as the purpose of HTA bodies in England and Germany is decreased by the dissimilarity in what factors take priority

in appraisals and to what end. This exemplifies how a seemingly scientifically objective concept such as *benefit appraisal* gets reinterpreted to fit national policy paradigms. This reinterpretation occurs when operationalising institutional goals during the process of normal decision-making.

## **5.6. Principles, Criteria and Methods of HTA Decision-Making**

The principles, criteria and methods of HTA decision-making are important elements of the policy paradigms in that they offer an insight into how the paradigms are operationalised. The principles of HTA decision-making are the values that reflect the policy paradigms at large whereas the criteria, and more importantly the methods of conducting HTAs, are the way these values get translated into practice. Before setting out these principles, criteria and methods, it is important to be aware of the distinction between different scientific methods for conducting HTAs, which all require slightly different evidential inputs and methods of analysis and are important to understanding the difference between the HTA paradigms in Germany and England. In a sense, the methods for conducting HTAs form the pool of potential instruments which decision-makers can choose from.

The most important inputs in HTA processes usually centre on evidence around clinical effectiveness, cost effectiveness, clinical benefits, side effects and other such measures. However, there are different methods to analyse the data that is presented in the context of such evidence. The most prominent methods for analysing the data include cost effectiveness (CEA) and cost-benefit (CBA) analyses. Sloan defines these as: “[...] formal methods for comparing the benefits and costs of a medical intervention in order to determine whether it is worth doing [...]” (Sloan, 1995, p. 3). Eichler, et al. (2010, p. 280) point out that there is no generally agreed definition of CEA, but it usually refers to the value for money.

In cost effectiveness analyses, which are used by NICE in England, the costs and benefits of a medicine or an intervention are compared with the current standard therapy for a given disease indication. Cost effectiveness analyses express clinical outcomes and benefits in non-monetary units such as quality-adjusted life years (QALYs) and costs in monetary units such as incremental costs, two concepts which will be explained in turn.

NICE defines QALYs as:

A measure of the state of health of a person or group in which the benefits, in terms of length of life, are adjusted to reflect the quality of life. One QALY is equal to 1 year of life in perfect health. QALYs are calculated by estimating the years of life remaining for a patient following a particular treatment or intervention and weighting each year with a quality of life score (on a zero to 1 scale). It is often measured in terms of the person's ability to perform the activities of daily life, freedom from pain and mental disturbance (NICE, 2014b).

In other words, QALYs are used to evaluate a pharmaceutical according to the increase in quality of life it provides and how much a patient's life might be prolonged as a result of treatment with the pharmaceutical. However, QALYs only express "the measure of the state of health of a person". In order for them to be useful in an HTA process that emphasises cost effectiveness, they have to be used alongside methods that express the marginal costs, additional costs or cost savings per QALY.

In England, the incremental cost effectiveness ratio (ICER) has been used to express the incremental costs associated with a treatment or intervention per QALY. In other words, expressing cost effectiveness as an ICER per QALY provides an indication of how much additional money will have to be spent on the new treatment in order to gain one extra QALY, i.e. one extra year in perfect health. However, once again the expression of an ICER per QALY is just one step in combining the two concepts in order to inform decision-making on cost effectiveness. In order to make sense of incremental cost per QALY measures, decision-makers need to know what would be considered a justified cost increase compared with an unjustified cost increase. Put differently, there needs to be system in place that allows decision-makers to understand *how much* cost increase is justified by *how much* clinical benefit. As we shall see, NICE employs cost effectiveness thresholds above which a new pharmaceutical product will usually not be considered a cost effective use of NHS resources.

The advantage of using ICERs and QALYs as measures of cost effectiveness is that they provide the decision-making agency with a comparative measure to evaluate cost effectiveness. QALYs can be used to measure the effect of a medicine on a person's health regardless of the illness for which that medicine is indicated. This means that clinical benefits and costs of a range of medicines can be measured across different disease areas and that, based on ICER per QALY measures, decision-makers are able to recommend the most cost effective medicine for coverage under a national health system. This might be considered important in a cash-limited health system

where opportunity costs – i.e. the microeconomic principle that every unit of money one spends in one area is not available to be spent in another – are an important consideration when providing health care. Connecting the ICER/QALY method with opportunity costs suggests that spending money on a treatment that is considered cost ineffective is not justified considering that benefits from alternative, and perhaps more cost effective treatments, will be foregone. While the use of QALYs is not uncontroversial and its critics argue that it does not allow for the consideration of important value judgements (Nord, 1999, p. 3-4)<sup>23</sup>, it represents an important methodological instrument that HTA decision-makers have employed.

Moving away from cost effectiveness analyses and on to an example of cost-benefit analysis, the so-called efficiency frontier is a relatively new method used for measuring and expressing the cost-benefit ratio of a given intervention. Thus far it is only used by IQWiG in Germany and while it is widely used in other economic evaluation contexts it is new in the context of measuring the cost-benefit ratio of health care interventions (Caro, et al. 2010, p. 1123). Drummond and Rutton describe the method in the following way:

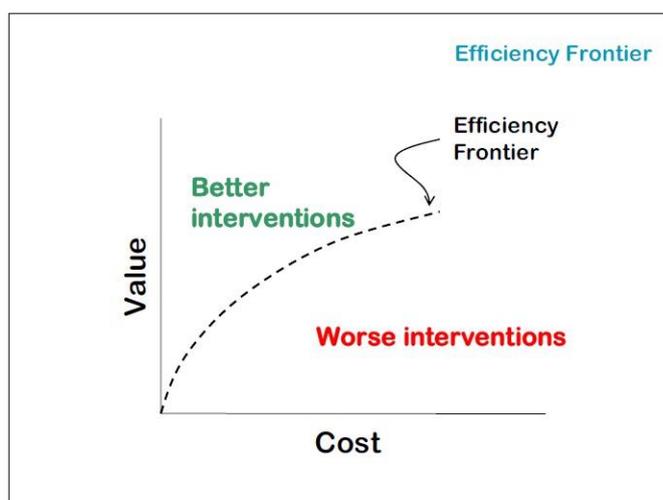
To compare costs and benefits in a particular therapeutic area [...] a diagram with costs on the X-axis and ‘value’ on the Y-axis [is constructed] and then [...] the existing therapies in this therapeutic area [are plotted] as points on the graph (Drummond and Rutten, 2008, p. 7),

with the most dominant existing interventions forming the “efficiency frontier” in the graph as expressed in their cost-benefit ratios. For products under assessment this results in a diagram such as the following:

---

<sup>23</sup> A detailed overview of the criticisms of QALYs and the associated ICER thresholds is beyond the scope of this thesis.

**Figure 4: The Efficiency Frontier**



(Source: IQWiG, 2008, p. 14)

Plotting different interventions on the graph aids decision-maker to judge which interventions provide the best benefits in comparison to the associated costs. While authors such as Caro, et al. (2010) argue that this method is useful, others such as Drummond and Rutten (Drummond and Rutten, 2008, p. 9) criticise it for the danger that it plots therapies on the graph whose cost effectiveness is not proven. This danger arises because the efficiency frontier demands that all medicines or interventions for a particular illness are included in the graph, even if effectiveness studies are not available for some of them due to their age (Drummond and Rutten, 2008, p. 9).

The brief overview of the methods for evaluating the benefits and costs of a given pharmaceutical product suggests that there is no such thing as *one* method for assessing and appraising a product's benefits. HTAs are complex processes in which a myriad of methods can be used. The choice of methods and the reasoning behind this need to be considered when examining what determines the outcome of pharmaceutical benefit assessments as they provide insights into the dominant HTA paradigm in a given country. More importantly though the empirical chapters in this thesis highlight that the way these methods or, to use Kuhn's (1962) words the 'rules of the game', operate in normal decision-making provides a deeper understanding of the effect that they have on the outcomes and on how a paradigm operates in practice.

## 5.7. England: Principles, Criteria and Methods of HTA Decision-Making

The principles that guide HTA decision-making at NICE are **consistency, equality and non-discrimination** (NICE, 2008). Of these principles consistency is one that is not mentioned as a guiding principle in any of the previously discussed tiers of the policy paradigm. Consistency can be viewed as an extension of and a reinforcing element of equality and non-discrimination. By ensuring that the decision-making process at NICE is comparable and that the same criteria are applied on a case-by-case basis a consistent approach promotes equality and non-discrimination. In order to apply the principle of consistency in practice, NICE has developed a ‘reference case’ for its assessments (NICE, 2008, p. 4). This reference case provides a framework of necessary criteria and methodological steps that technology appraisals at NICE must follow. It specifies aspects of HTAs including, but not limited to, what the relevant comparator should be (i.e. therapies routinely used in the NHS), how health effects are to be measured and expressed (i.e. in quality-adjusted life years (QALYs)) and what type of economic evaluations should be used (i.e. cost effectiveness analysis) (NICE, 2008, p. 30).

The principle of consistency is therefore put into practice by means of a reference case that stakeholders, including pharmaceutical manufacturers and decision-makers at NICE, can use as a roadmap when navigating through the pharmaceutical benefit assessment process. NICE explains the reasoning for using a reference case as a tool in the decision-making process in the following way:

The Institute has to make decisions across different technologies and disease areas. It is, therefore, crucial that the analyses of clinical and cost effectiveness undertaken to inform the appraisal adopt a consistent approach. To allow this, the Institute has defined a ‘reference case’ that specifies the methods considered [...] to be the most appropriate for the Appraisal Committee’s purpose and consistent with an NHS objective of maximising health gain from limited resources (NICE, 2008, p. 31).

This reasoning is congruent with the wider paradigm under which the NHS operates in that it makes reference to the limits of health care resources. It therefore becomes necessary to make decisions across disease areas because the English paradigm accepts that resources are limited. Therefore, if one is to uphold two other important paradigmatic features of the health care system in England, i.e. equality and non-

discrimination, the best way to do so is by ensuring a consistent approach in decision-making, thereby reflecting a procedural value akin to a *fair process*. Following on from this, the methods laid out in NICE's reference case reflect the paradigmatic need to make decisions across disease areas whilst ensuring a fair process by stipulating that the reference case needs to be adhered to.

The principle of consistency, and by extension of quality and non-discrimination, is operationalised by employing NICE's reference case. The reference case is built on a conceptualisation of criteria for HTA decision-making such **clinical and cost effectiveness (across disease areas)**. According to NICE,

[...] technologies can be considered clinically effective if, in normal clinical practice, they confer an overall health benefit, taking into account any harmful effects when compared with other relevant treatment alternatives. Technologies can be considered cost effective if their health benefits are greater than the opportunity costs measured in terms of the health benefits associated with programmes that may be displaced to fund the new technology. In other words, the general consequences for the wider group of patients in the NHS are considered alongside the effects for those patients who may directly benefit from the technology of interest (NICE, 2008, p. 9).

In addition to providing the definitions of clinical and cost effectiveness in the context of NICE's decision-making process, the above excerpt from NICE's guide to technology appraisals illustrates that the needs of the many, i.e. the "wider group of patients in the NHS" (NICE, 2008, p. 9) are balanced with the needs of the few, i.e. "the patients who may directly benefit from the technology of interest" (NICE, 2008, p. 9). It also suggests that clinical effectiveness is a necessary but not a sufficient condition for a positive appraisal by NICE.

The methods and criteria of NICE's decision-making process are designed to accommodate the paradigmatic idea that a balance has to be struck between different patients' needs. Health benefits are expressed in terms of QALYs. This method allows for a comparison between different pharmaceutical products, for different disease areas, in order to assess their comparative effectiveness in relation to their costs. In addition, the incremental cost of an intervention is calculated and compared with the additional clinical benefit as expressed in QALYs. This gives rise to the ICER, which compares the mean costs with the mean clinical outcomes of a treatment (NICE, 2014b).

As long as thresholds are defined, employing ICERs allows decision-makers to decide which interventions are cost effective and which ones are not. The underlying assumption of the use of ICERs is that there are thresholds above which the additional costs incurred are no longer justified by the extent of expected additional clinical benefits. However, the way in which these thresholds are set (NICE's current threshold is at £30,000 per QALY) is contested. Further elaboration on this is beyond the scope of this thesis. In the context of this thesis, it is important to highlight that ICERs and QALYs are more than scientific tools for conducting HTAs. They are the logical methodological choice that emerges from a paradigm that a) accepts that health care resources are limited and b) requires balancing the needs of the many with the needs of the few. Williams summarises this choice in the following words:

The methodological underpinnings for TAs [technology appraisals] derive from non-welfarist health economics, and are driven by the objective of maximizing population health subject to budget constraint. This leads to the use of a decision criterion formulated in terms of the extra costs that would have to be incurred in order to bring about a unit improvement in health. [...] But ranking technologies according to their cost-per-QALY score still does not get us to our destination, because we still lack a cut-off point beyond which we say “this far but no further” [...] (Williams, 2004, pp. 6-7).

In addition to clinical and cost effectiveness as criteria for decision-making, the Appraisal Committees (ACs) at NICE are to take into consideration the **clinical need** of patients, “the potential for long-term benefits to the NHS of **innovation**” (NICE, 2008, p. 52) as well as NICE's guideline on **social value judgments** (NICE, 2008a). More recently, NICE has also developed a supplementary advice on appraising **end-of-life treatments** that permits ACs to recommend a pharmaceutical product even if its ICER is above the usually accepted threshold of £30,000 per QALY if the product is licensed for the end-of-life treatment of a terminal disease (NICE, 2009a). This was developed as a result of a number of appraisal processes, especially on expensive end-of-life cancer treatments, in which ACs were not able to recommend the use of the treatments on clinical and cost effectiveness alone, but needed additional decision-making criteria for the special situations pertaining to the end-of-life stages of diseases such as cancer. Thus, while clinical and cost effectiveness are the most important pillars of the HTA decision-making processes at NICE, the HTA paradigm has developed in such a way

that permits for a consideration of other factors and values in special circumstances such as end-of-life settings or innovative developments.

In summary, the principles and criteria of HTA decision-making as they are laid out in NICE's guide to technology appraisals are a reflection and extension of the values found in the wider health care context. **Equality and non-discrimination** are expressed in the principle of **consistency** and operationalised by means of NICE's **reference case**. This ensures a fair process and provides stakeholders and decision-makers with a decision-making algorithm in that they know what steps need to be followed and what criteria need to be met. In the case of cost effectiveness the algorithm is clear in that products with ICERs above £30,000 per QALY need to meet other criteria such as the end-of-life criteria or innovation considerations in order to be recommended for use (NICE, 2008). In this way, while the English HTA decision-making paradigm provides a consistent algorithm, it also allows exceptions if the algorithm does not do justice to the benefits of the treatment in question.

What is striking about the English HTA paradigm, especially in comparison to the German paradigm explored in the next section, is the unequivocal acceptance of the limits of health care resources. The decision-making structures are designed in a way that permits fair decision-making whilst knowing that the NHS cannot provide everything that might be desirable or even necessary. This unequivocal acceptance leads to a HTA paradigm that seeks to strike a balance between the needs of the many with the needs of the few, whilst providing ways around it in special circumstances.

## **5.8. Germany: Principles, Criteria and Methods of HTA Decision-Making**

Like the English HTA decision-making paradigm, the principles that guide HTA decisions in Germany are a reflection of the wider health care context. The FJC's Rules of Procedure include the principles of **transparency and legal soundness, the generally accepted state of medical knowledge**, the need for considering justified **interests of stakeholders and efficiency** (G-BA, 2013). The principle of legal soundness has been elaborated on in a previous section of this chapter and reflects the fact that appeals opportunities other than taking legal action are limited in the German HTA context. The 'generally accepted state of medical knowledge' reflects the inclusion of this value in Article 1 of the Social Code Book V as does the principle of efficiency in Article 12 of the Social Code Book V (BMJV, 2013). Thus, except for the

specification of considering issues of quality and health care provision when making decisions, the only newly introduced principle in the FJC's Rules of Procedures is the need to consider the interests of stakeholders; how this principle is operationalised is addressed in the empirical chapters that follow. IQWiG as the evidence review body, or assessment body, in the German context highlights its **independence** as a scientific body, **transparency** and the principle of **evidence-based medicine** as its guiding principles when making decisions (IQWiG, 2011a).

However, the most meaningful insights into HTA decision-making in Germany cannot be found in the above principles, but in the way the most important decision-making criterion, namely that of '**therapeutic benefit**', is conceptualised and operationalised. Therapeutic benefit is understood as the additional benefit of the appraised medicine in comparison to currently available alternatives. The Ordinance for Assessing the Benefits of Pharmaceuticals specifies that:

the benefit of a pharmaceutical is the patient-relevant therapeutic effect, in particular in respect of the improvement in the state of health, the reduction of the duration of the disease, longer survival, the reduction of side-effects or an improvement in quality of life (Bundesgesetzblatt, 2010).

This conceptualisation of therapeutic benefit as an expression of *patient relevance* suggests that the clinical effectiveness of a pharmaceutical product, i.e. its ability to achieve what it sets out to do in a normal clinical setting, is not sufficient for it to receive a positive appraisal in the form of the assignment to a high additional benefit category. Rather, clinical effectiveness in Germany is conceptualised and operationalised as a concept of categorical quality which does not automatically amount to an additional benefit. IQWiG specifies that the concept of patient relevance includes how a patient feels, how he/she can go about his/her daily activities, whether a patient survives as well as questions of mortality, morbidity and health-related quality of life (IQWiG, 2011a). In the German context all clinical trial endpoints have to be patient relevant in order to be included as valid endpoints when making appraisal decisions. The concept of patient relevance can be considered the most important criterion in HTA decision-making.

Despite patient relevance being the most important criterion of HTA decision-making, it is not clear which principle or value of the wider health care system or the HTA process it reflects or arises from. This is different from the English paradigm in

which the connections between the different paradigmatic value tiers are more easily discernible. Moreover, the empirical analyses in the next chapters show that, for example in the case of Boceprevir (G-BA, 2012) and Telaprevir (G-BA, 2012a), the operationalisation of patient relevance does not always sit easily with the generally accepted state of medical knowledge. In the case of Boceprevir and Telaprevir, both indicated for chronic hepatitis C of genotype 1, this meant that the main clinical endpoint in the form of sustained virological response (SVR) was initially not accepted by IQWiG as being patient relevant even though medical experts agreed that it was the most important indicator of successful treatment in hepatitis C (G-BA, 2012; G-BA, 2012a). The case illustrates that it is not clear *how* patient relevance should be defined on a case-by-case basis and which methods should be used to demonstrate patient relevance if the methods employed in demonstrating clinical effectiveness are not deemed satisfactory by IQWiG or the FJC. Germany's HTA decision-making paradigm includes a conceptualisation of clinical effectiveness as the patient relevant therapeutic effect, but in comparison to England it does not provide a decision-making construct such as NICE's reference case to offer guidance on the operationalisation of patient relevance.

Despite the lack of a reference case and despite its emphasis on evidence-based medicine, the importance of patient relevance is congruent with a paradigm that, as a result of the Nikolaus-case court ruling (Das Bundesverfassungsgericht, 2005), also needs to cater for the right to treatment even when effectiveness evidence is not available or sufficient. The concept of patient relevance, i.e. what really matters to the patient, may be seen as an attempt to accommodate this unique HTA paradigmatic feature. It may also explain why, apart from the principle of efficiency, matters of costs or cost effectiveness do not feature as decision-making criteria in German pharmaceutical benefit assessments. What matters solely is the additional benefit of a new medicine in comparison to what is currently available, not whether this medicine is cheaper or more expensive in relation to the benefits it offers. Additionally, a cost-benefit assessment, in which IQWiG would conduct an economic evaluation using the 'efficiency frontier' method described earlier, is not conducted automatically but has to be commissioned separately from an early benefit assessment. In contrast to England, where there is an acknowledgement of the limits of health care resources, in Germany the opposite appears to be the case in that there is a reluctance to engage with questions of costs and cost effectiveness in pharmaceutical benefit assessment processes. This

claim is substantiated by the empirical case studies presented in the next chapters in that cost effectiveness arguments raised by the pharmaceutical industry and other stakeholders are discarded as irrelevant by the FJC.

In addition to the importance of patient relevance and the comparative lack of a consideration of cost issues, the benefit assessment within disease areas rather than across disease areas is an important feature of the German HTA paradigm. In a way this is a result of the decision to conduct benefit analyses rather than cost-benefit analyses in the German context which in turn is a result of a specific paradigmatic idea of what purpose HTAs should serve. In other words, in the German context HTAs are supposed to provide insight into the additional benefits of one medicine in comparison to another in the same disease area rather than insight into which medicine can be considered cost effective across disease areas. As IQWiG points out this decision is based on a value judgment about whether it is right or fair to compare medicines for different diseases and potentially different disease severity with one another (IQWiG, 2011a). Within the German decision-making paradigm it is not deemed acceptable to assess the value of benefits across disease areas and both the FJC and IQWiG are explicit about this.

To summarise, there are three features that stand out from the principles and criteria of HTA decision-making in Germany, namely the conceptualisation of **therapeutic benefit as patient relevant effect**, the comparative **lack of the consideration of cost issues** and the **benefit analyses within disease areas** rather than across different diseases. While the comparative disregard of cost issues is a reflection of the institutional purpose of the German HTA paradigm, i.e. its purpose is to inform price setting, the conceptualisation of therapeutic benefit and the decision to assess benefit within disease areas reflect a specific value view of what is or is not acceptable when making decisions in health care.

It is worth noting that the principle of innovation is not included as a criterion in the methodological guidelines. This may be because the highest benefit category is thought of as a recognition of the innovative character of a new medicine. However, considering that the early benefit assessment process was introduced to mitigate against an increasing number of so-called pharmaceutical innovations that were, on closer inspection, not innovative in comparison to already existing medicines, it is interesting that innovation does not constitute a separate criterion. The empirical analysis in the proceeding chapters sheds light on whether innovation is considered in normal HTA decision-making processes in Germany.

What is striking about the German HTA paradigm is the idea that clinical effectiveness alone is not a sufficient indicator to prove that a new medicine provides an additional benefit. What is exceptional about the German paradigm is the criterion used to assess comparative effectiveness, namely that of patient relevance. While the common approach to evaluating a pharmaceutical product's comparative effectiveness is to analyse its statistically significant parameters and to put these in relation to costs, Germany's approach introduces an additional hurdle that clinical effectiveness needs to meet in that it has to be *patient relevant*. However, in doing so the methodological guidelines provide little insight into how patient relevance should be defined and whether there is a decision-making algorithm that indicates the thresholds or minimum criteria for the assignment of a new medicine to one of the six possible benefit categories<sup>24</sup>. The latter might change in the future as IQWiG is currently developing new methodological guidelines. However, for the cases that are analysed in this thesis the lack of clear guidance on the operationalisation of patient relevance and the benefit categories was one of the points raised in each assessment process.

Finally, when examining the theoretical and the empirical aspects of the German HTA paradigm a dissonance arises that is difficult to disentangle. When analysing the methodological guidelines and legislative framework one faces a paradigm that is characterised by cautious and seemingly conscious attempts to avoid some of the ethical conundrums of health care decision-making. That is to say that the HTA paradigm is set out in such a way that even the assignment of a medicine to a negative benefit category will not automatically lead to its exclusion in the health care benefit basket. Similarly, the decision to appraise benefits within rather than across disease areas means that one group of patients will never gain benefits at the expense of others. In times of health care rationing and prioritisation these are striking paradigmatic features. However, when it comes to the operationalisation of the paradigm, and especially the outcomes of benefit assessment processes in different cases, the FJC's appraisal of the benefits of a new product is frequently more restrictive than NICE's. Thus arises a dissonance within the paradigm in that it contains ideas that are more favourable to individual patients' needs than NICE's, but it is lacking on guidance on how to operationalise these ideas so that theory and practice meet.

---

<sup>24</sup> See previous chapter for an overview of the benefit categories.

## **5.9. The Case of Apixaban: How are HTA Paradigms Operationalised?**

In this section I introduce the case of Apixaban (G-BA, 2012b; NICE, 2012a) as an example of how HTA paradigms are operationalised in England and Germany. While the cases discussed in the next chapters are cases in which challenging questions around the evidence arose, the case of Apixaban illustrates a comparatively straightforward example of the reasoning that led to the final outcome during the normal HTA decision-making process. That is to say that the previously presented decision-making criteria, especially in relation to the methods used, were applied without much controversy by NICE, IQWiG and the FJC respectively. Yet, despite a lack of controversy within the decision-making institutions and despite the consideration of the same pieces of evidence, the institutions came to slightly different decisions on the benefit of Apixaban. As a result, Apixaban is a useful case to illustrate how differences in the rules of evidence can lead to slightly different outcomes even in the seemingly most straightforward cases of pharmaceutical benefit assessments.

Table 5.2. provides an overview of the product, the dependent variable as well as the main issues that were considered by NICE, the FJC and IQWiG in the decision-making process on Apixaban.

**TABLE 5.2. – Case Study: Apixaban**

(Indicated for: Prevention of thromboembolic events (i.e. blood clots) after total hip or knee replacements)

	NICE (NICE, 2012a)	FJC (G-BA, 2012b)
<i>Dependent variable: Outcome of benefit assessment</i>	Recommended	Different benefit appraisal for the two patient populations: <ul style="list-style-type: none"> <li>- For knee-replacement population: No additional benefit</li> <li>- For hip replacement population: Marginal additional benefit</li> </ul>
<i>Reasoning/discussions/topics raised/public context</i>	<ul style="list-style-type: none"> <li>- More clinically effective and cheaper than at least one comparator (Enoxaparin)</li> <li>- Two (ADVANCE 2 &amp; 3) out of four good quality RCTs comparing Apixaban with Enoxaparin relevant to UK clinical practice. Primary efficacy endpoints was a composite of all incidences of venous thromboembolisms (VTE), i.e. pulmonary embolisms, symptomatic and asymptomatic deep vein thrombosis (DVT)</li> <li>- Issues with trial outcome components: Clinical specialists highlighted that there were limited data to show relationship between one of the major components of the composite primary outcome, i.e. the relationship between asymptomatic DVT and clinically relevant VTE. Appraisal Committee agreed that it was a widely used outcome measure and was relevant for consideration.</li> <li>- Advantage of oral application</li> </ul>	<ul style="list-style-type: none"> <li>- Separate consideration for the two patient populations for which Apixaban is licensed because surgery methods and the associated risks are different for knee and hip replacements.</li> <li>- Two relevant RCTs (ADVANCE 2 &amp; 3)</li> <li>- Issues with trial outcome components: Asymptomatic deep vein thrombosis (DVT) is not patient relevant because in routine clinical practice it does not get treated. Therefore, endpoints that include DVT cannot be considered. Correlation between asymptomatic DVT and other clinically relevant venous thromboembolisms (VTEs) not demonstrated in the scientific data.</li> </ul>

<p><i>Differences/similarities</i></p>	<p>→ <b>Positive recommendation</b></p> <ul style="list-style-type: none"> <li>- Accepted 2 RCTs</li> <li>- Questioned correlation between asymptomatic DVT and clinically relevant VTEs, but accepted clinical outcomes presented by manufacturer as widely used outcome measures</li> <li>- Positive recommendations even though ICERs were not clear but all likely under the threshold</li> <li>- Advantages of oral treatment mentioned</li> </ul>	<p>→ <b>Positive &amp; negative recommendation for one of the two patient populations</b></p> <ul style="list-style-type: none"> <li>- Accepted 2 RCTs</li> <li>- Questioned correlation between asymptomatic DVT and VTEs, especially in relation to its patient relevance. Did not consider asymptomatic DVTs as a patient relevant endpoint.</li> <li>- Cost effectiveness not considered</li> <li>- Advantages of oral treatment not mentioned</li> <li>- Oral hearing: Extensive discussions around the patient relevance of endpoints</li> </ul>
--	---	--

## 5.10. Discussion

Table 5.2. shows that the outcome of the benefit assessment process on Apixaban was positive in Germany and England. However, the FJC qualified its positive appraisal in the sense that it only found an additional benefit for one of the two patient populations for which Apixaban is indicated, namely that of the hip replacement population (G-BA, 2012b). The FJC and IQWiG came to this conclusion because the surgery methods and associated risk for venous thromboembolisms (VTE) and deep vein thrombosis (DVT) are different for the two patient populations, hence they looked at the two populations separately and found that the statistical results of the clinical trials were different for the two population (G-BA, 2012b).

It is interesting that in the case of Apixaban both NICE and the FJC accepted the same evidence in the form of two RCTs as being reflective of routine clinical practice (G-BA, 2012b; NICE, 2012a). As will be seen in later chapters, this agreement on the evidence base for a benefit assessment is not always the case. The records also show that the question of whether one the main components of the composite primary outcome, i.e. whether asymptomatic DVT is relevant to the occurrence of VTE, was considered by both NICE and the FJC (G-BA, 2012b; NICE, 2012a). After hearing different opinions from medical experts on this issue, NICE decided to accept the component as part of the composite primary outcome of the clinical trial.

The FJC heard similar contrasting opinions by medical experts, but in contrast to NICE it decided that, based on the criterion of patient relevance, asymptomatic DVT would not be considered in this appraisal (G-BA, 2012b). The FJC argued that it was not a patient relevant endpoint because in routine clinical practice patients would not get treated for an asymptomatic DVT (G-BA, 2012b, p. 6). The reasoning behind this was that it is only in the context of a clinical trial that one would diagnose an asymptomatic DVT and treat it, which means that it is not relevant to the wider patient population. Furthermore, IQWiG and NICE were not convinced that there is sufficient scientific data to suggest a correlation between asymptomatic DVT and clinically relevant VTEs. The case of Apixaban thus shows how HTA bodies can come to different conclusions on the same problem when the respective decision-making paradigm is applied.

Theoretically the different appraisal of the patient populations in Germany could mean that sickness funds could negotiate a lower price for the knee replacement patient population because no additional benefit was found for this population. Whether this

would be feasible and whether it is actually done in practice is a difficult question to answer as the price negotiations between the sickness funds and pharmaceutical manufacturers are confidential, meaning that the public and researchers do not have access to the records of these meetings. However, empirical evidence gained as part of the interview process suggests that the price is aggregated for all patient populations rather than agreeing on a different price for different populations (Interviewee No. 23, 2013, p. 7). A further implication of the FJC's decision in the case of Apixaban could be that the product does not find its way into routine clinical practice because of the lack of additional benefit. However, while opening interesting empirical questions, this question is beyond the scope of this thesis.

The case of Apixaban is a good example of how HTA paradigms are applied because the clinical evidence on it was comparatively straightforward, apart from the above mentioned issue of the correlation between asymptomatic DVT and VTEs. Moreover, outside pressures such as media involvement or patient group action were virtually non-existent in this case because it was uncontroversial. The costs of the use of Apixaban were also comparatively low. Apixaban is a case in which factors such as clinical evidence, external pressures and cost considerations are controlled for. Yet, despite controlling for these variables we observe a slightly different outcome on the dependent variable in this case. The reason for this cannot be found in clinical and cost effectiveness issues per se but rather in the interpretation thereof, i.e. in how the HTA paradigms are operationalised. The application of a patient relevant paradigm in Germany led to a slightly more restrictive overall appraisal of Apixaban while the application of the cost effectiveness paradigm in England led to a positive recommendation as Apixaban was more clinically effective and cheaper than at least one of comparators (NICE, 2012a). In other words, while the assessment of the weaknesses of the evidence was similar in England and in Germany, the appraisal of these weaknesses was different due to differing decision-making criteria. In the end, this led to a slightly more generous overall appraisal by NICE than by the FJC.

## 5.11. Conclusion

Before turning to a more in-depth analysis of how the English and German HTA paradigms are operationalised in practice, I offer a number of concluding remarks about what the previous discussion suggests about what we might expect about HTA paradigms in operation. This is particularly relevant to the policy implications of HTA decision-making in that how paradigms are operationalised plays a role in whether health care funding decisions are deemed to be generous or restrictive.

The differences between the English and the German HTA paradigms are multiple. NICE operates on a paradigm that emphasises **consistency, equality and non-discrimination**. It operationalises these values by employing a ‘**reference case**’ that includes clinical and cost effectiveness, innovation, social value judgments and end-of-life considerations as decision-making criteria (NICE, 2008). It subscribes to the idea that costs matter and that a balance needs to be struck between the needs of patients with different conditions. To this end, NICE’s paradigm contains decision-making algorithms and thresholds, most importantly in the form of **ICERs**, that provide guidance for stakeholders and decision-makers alike. ICERs allow for the comparison **across disease categories** in order to identify the most cost effective treatments across different disease indications and interventions. Moreover, the English paradigm specifically acknowledges that health care resources are finite and that this calls for transparent, fair and accountable decision-making.

By contrast, the FJC and IQWiG operate on a paradigm that puts the concept of **patient relevance** at the heart of decision-making. This reflects a wider health care framework that emphasises solidarity and the right to health care treatment even when evidence is not conclusive or not available. Questions of costs and the balance between different patients’ needs are not considered, at least there is no sign of it in the statutory framework or methodological guidelines. In contrast to NICE, the FJC only compares clinical benefit categories **within disease categories**. The underlying rationale for this appears to be the belief that it would not be fair or solidaristic to compare interventions for patient populations with very different diseases.

Most importantly, NICE and the FJC/IQWiG operate on different understandings of clinical effectiveness. In NICE’s case clinical effectiveness gains meaning in relation to costs through employing the concept of the ICER. For IQWiG and the FJC it gains meaning in relation to the idea of patient relevance. In light of this,

the concept of thresholds is put to different uses. At NICE, thresholds, or ICERs, are used to indicate the maximum incremental cost of a medicine that NICE considers an appropriate use of NHS resources in relation to expected clinical benefits. At the FJC and IQWiG the use of thresholds is less formalised in that the operationalisation of the six benefit categories is not always clear. However, the operationalisation of benefit by means of categories suggests that there are thresholds for the maximum price that is justified by the expected additional clinical benefit.

Despite the important distinction between how, and to what end, thresholds are operationalised in England and Germany the underlying rationale for conducting HTAs appears to be similar. It suggests that there is a limit to the price that a society is willing to pay for pharmaceutical products and that this limit is dependent on the additional clinical benefits that a product provides. Thresholds play an important role in this regard as they are a methodological expression of the aforementioned limit. This has important implications for the framework of policy paradigms. In the case of pharmaceutical benefit assessments policy-makers established that there should be limits to what the health care system pays for a new product and that these limits should be related to the clinical benefits. However, neither in England nor in Germany did policy-makers define the nature of these limits, how they should be operationalised and where they should be set. The articulation of these specifics occurs when the paradigm is established in normal decision-making, i.e. when decision-makers conduct HTAs. Normal decision-making thus serves the purpose of articulating the specifics of the paradigm and resolving tensions that it gives rise to, such as the ones discussed in section 5.4. The question of what this means in practice is presented in the next chapters. For now, suffice it to conclude that the underlying idea of the HTA paradigm in England and Germany is that costs, or prices, need to be justified by clinical benefits, but this underlying idea is translated into practice by different rules and methods, thereby affecting outcomes.

In comparison to the FJC, which focuses on the criterion of patient relevance, NICE's decision-making paradigm allows the consideration of a variety of issues and criteria. The question arises whether this leads to more generous outcomes of pharmaceutical benefit assessments. Arguably, the possibility of considering multiple issues when making decisions would lead to more generous decisions as decision-makers have a wider array of argumentative rationales to employ when making decisions. However, as the next chapters show, this logic does not hold true when

looking at the empirics of individual cases. This is mainly because of the apparent existence of a value core in which cost effectiveness trumps other considerations. While the FJC does, overall, take more negative decisions on the additional benefit of a new pharmaceutical product, the net result of this has little impact as it only impacts on price negotiations and not on the availability of the product.

There are two further preliminary conclusions I draw from the previous discussion. The first one is that both theory and empirics suggest that there is a core of values within HTA paradigms, meaning that some concepts are weighted more than others. For the English paradigm this appears to be the principle of cost effectiveness while for the German paradigm it is the principle of patient relevance. The idea of a value core, whilst having empirical implications, is especially relevant with regards to the theoretical contribution this thesis makes in that it confirms that paradigms impact on what is considered possible in a policy area. The identification of the relevant values, and their relative importance, might help in better understanding this impact.

The second additional conclusion relates to the question of whether ideational or institutional variables matter when it comes to the outcome of pharmaceutical benefit assessment processes. The above analysis and the example of Apixaban show that ideational factors such as conceptualisations of clinical effectiveness and patient relevance play a large role when determining the outcome of pharmaceutical benefit assessments. However, the extent to which this role matters when it comes to the availability of new medicines depends on the institutional purpose of HTAs. This explains why a pharmaceutical product is usually still available in Germany even if the FJC's decision on it was negative because HTAs are used to inform price setting. The institutional construct in England is different in that a negative recommendation by NICE usually means that a product is not made available for routine use in the NHS. Interestingly though, a clinical commissioning group (CCG) could still decide to fund the respective product even if its use is not recommended. However, the financial constraints within the NHS are such that this is unlikely to happen. This shows that there might be situations in which a decision is *institutionally possible*, but politically or financially not feasible due to external constraints.

In addition to a close interconnection between ideational and institutional factors that determine the outcome of pharmaceutical benefit assessments, the policy implications emerging from ideational factors appear to be kept in check by an institutional construct that overrides the ideational one. This holds true for both England

and Germany. However, this should not lead to the conclusion that institutional variables are more significant than ideational ones in determining the outcome of pharmaceutical benefit assessments. As seen in this chapter, institutional constructs are a reflection of ideational frameworks in the form of the values that underlie the wider health care systems. The interplay between ideational and institutional variables is varied and complex. At times they mutually reinforce each other, they can provide checks and balances to the other and they can act as overriding principles in some cases. This complexity is explored in the following chapters to shed further light on what determines the outcome of pharmaceutical benefit assessments in health care systems that employ formalised HTA procedures.

## Chapter 6

### **Policy Paradigms in Operation I: The Conceptualisation of Evidence**

#### **6.0. Introduction**

This chapter marks the first of three chapters that present the empirical analysis of what determines the outcome of pharmaceutical benefit assessments in countries that employ formalised health technology assessment (HTA) processes. As such it is the beginning of the analysis of how HTA paradigms operate in practice, that is how they are applied, articulated and established in ‘normal’ decision-making processes. The chapter commences by summarising the themes that emerged from the data. Six themes emerged in total (table 6.1.) and they are presented in order of prevalence. The summary of the themes is followed by an in-depth analysis of case studies of pharmaceutical benefit assessments in order to demonstrate how the findings arise from the empirical data. At the end of every case study analysis an overview of the main research findings is provided. Overall, the empirical evidence suggests that the outcome of pharmaceutical benefit assessments is largely determined by the conceptualisation and interpretation of scientific evidence, which in turn is shaped by the HTA policy paradigm in a given health care context, thus confirming the assumption that ideational approaches can contribute to explaining the outcomes of pharmaceutical benefit assessments.

Following the summary of the themes that emerged during the data analysis process, the first part of the empirical analysis is presented in this chapter. The main focus is theme one (table 6.1.) as it was raised in all of the cases. It is the question of what constitutes appropriate evidence and how its weaknesses and strengths should be assessed. The chapter demonstrates that the way strengths and weaknesses in the evidence are addressed by HTA bodies, and indeed what constitutes evidence in the first place, is connected to different ‘rules of evidence’ (Majone, 1989). These rules give an idea of what is considered relevant in a given context, thus supporting Cartwright’s and Hardie’s (2012) assertion that pieces of evidence can only support decision-making if they are relevant to the problems that decision-makers and policy-makers face.

Empirically, the analysis of the cases of Cabazitaxel, Eribulin and Ipilimumab form the centre of this chapter. They represent cases in which NICE (2012b; 2012c; 2014c), the FJC (G-BA, 2012c; 2012d; 2012e) and IQWiG (2012; 2012a; 2012b) accepted the same evidence in the form of randomised controlled trials (RCTs). This led to a similar outcome in the dependent variable despite applying different HTA decision-making criteria to the evidence. The chapter traces the substantive arguments of the decision-making processes in the above cases in order to compare and contrast how the final outcome of the benefit assessments was arrived at.

As outlined in chapter 3, the data was analysed by employing process-tracing methods. The process-tracing exercise focused on tracing the reasoning processes in the ten embedded case studies, that is on identifying the issues, values and ideas that were employed by decision-makers to formulate the outcome of the benefit assessments. This explains the relative neglect of structural elements of the decision-making processes in the empirical analysis that follows. That is to say that, institutionally and structurally, the step-by-step HTA processes (see chapter 4) were consistent amongst the cases. Unless otherwise stated in the analysis, the reader can assume that the decision-makers followed the institutionally laid out procedural stages for conducting pharmaceutical benefit assessments. This consistency permitted focusing on the decision-making reasoning in order to address the research question.

In summary, this chapter introduces the themes that emerged during the analysis. It then discusses the first theme (see table 6.1.) by illustrating what constitutes evidence in Germany and England by reference to the cases of Cabazitaxel, Eribulin and Ipilimumab. The discussion highlights that the ultimate decision by NICE and the FJC was based on the same scientific evidence. However, different decision-making criteria, or rules of evidence (Majone, 1989), were applied to the evidence, which in turn meant that *different reasons for similar decisions* were given. In terms of using paradigms to explain outcomes of pharmaceutical benefit assessments this is a significant finding in that it suggests that different paradigms do not, as a matter of course, lead to different outcomes.

The chapter also includes the discussion of theme five (see table 6.1.), namely the suitability of paradigms for ‘special’ cases such as chronic diseases. This theme is discussed here because the questions that were raised fit with the wider questions of what constitutes evidence and how it should be assessed in relation to specific diseases. Even though theme five was not as prevalent as the other themes because it only

pertained to the cases of chronic diseases, it gives informative insights on the potential limits of the English and German HTA paradigms as identified by stakeholders. Its discussion therefore highlights the extent to which HTA paradigms are applied flexibly during normal decision-making where ‘special’ cases are concerned.

## 6.1. Empirical Results

The empirical research that forms the basis for this study exemplifies the challenges and controversies that arise from employing evidence-based measures to inform decision-making in the health care context. Having examined how decisions were reached in ten cases of new pharmaceutical products that were assessed and appraised by NICE, the FJC and IQWiG, one of the conclusions I draw is that the way concepts of evidence and data are understood and interpreted contributes greatly to how decisions on benefit assessments of products are made. The term ‘interpretation’ refers to the kind of criteria that is applied in the decision-making process. Discussions about the appropriateness of data and the available evidence feature prominently in the publicly available assessment and appraisal documents on the ten cases. The scientific base and interpretation thereof was also a common theme that was raised in the interviews that I conducted with stakeholders in both Germany and England. It is the most prevalent theme that emerges from the data and its properties and questions are outlined under theme one in table 6.1.

Table 6.1.<sup>25</sup> provides an overview of the six themes raised in the pharmaceutical benefit assessment documents and stakeholder interviews. The way these themes were presented in the respective documents and the interviews allows inferences about what determines the outcome of pharmaceutical benefit assessments. The table shows that most themes that were raised in the relevant documents related to questions around evidence and technical problems such as whether the evidence included the right choice of comparator. As we shall see in this chapter, theme one relates to the permissibility, quality and validity of evidence. The first question that NICE, IQWiG and the FJC

---

<sup>25</sup> The columns on England and Germany in table 6.1. illustrate the focus and direction in which the themes were addressed by the respective institutions. Questions that were primarily raised by stakeholders such as physicians’ professional bodies are highlighted as such in the relevant columns. Where such highlights are lacking, it indicates that these questions were considered by most stakeholders, including decision-makers, rather than being limited to certain stakeholders.

engage with during a HTA process is the question of which evidence to consider. Questions around the quality and validity of the presented evidence and whether or not the evidence meets the respective decision-making criteria form a substantial part of the benefit assessment processes. The empirical data suggests that this holds true for England and Germany. However, the way questions are addressed and resolved differs in the two countries, which leads to the consideration of different evidence and helps explain differing assessment outcomes in some cases. In the case of Retigabine (table 6.6.), for example, NICE appraised all of the evidence submitted by the pharmaceutical manufacturer whilst IQWiG and the FJC accepted none of the evidence on the basis that none of the presented studies compared Retigabine with the appropriate comparator.

Themes two, three and four – the choice of comparator product, the patient population subgroup divisions and the operationalisation of criteria for HTA decision-making and the role of algorithms - could have been discussed as sub-themes of theme number one as they are all related to evidential questions. These evidential questions focused on whether appropriate comparators (theme two) and subgroup populations (theme three) were identified in the evidence and how study results on these should be interpreted (theme four). They constituted distinct and in-depth discussions in the appraisal documents and their importance was also highlighted by stakeholders who were interviewed. This warrants for a discussion of themes two-four in their own right, which is presented in chapter 7.

Theme five (the suitability of paradigms for ‘special’ cases such as chronic diseases) and theme six (political power and pressures) did not feature in every case, but only in the cases of Retigabine and Fingolimod (table 7.1.), and Abiraterone (table 8.1.) and Fingolimod respectively. However, they raised a sufficient number of noteworthy points about paradigms that merit their discussion as separate themes in this chapter and chapter 8. Whilst themes one-four are similar in that they centre on evidential questions, the emergence of themes five and six seems to suggest that there are cases in which the dominant, or the emerging dominant, HTA paradigm might be challenged, or might not provide satisfactory answers in the light of the nature of some cases. For example, in the case of Retigabine and Fingolimod the suitability of the dominant paradigm was called into question because of the chronic nature of epilepsy and multiple sclerosis respectively, which, according to some stakeholders, means that the benefits of these drugs should not be assessed by the same standards as non-chronic conditions.

Despite the differences in the institutional design of evidence-based measures to inform pharmaceutical coverage decisions in England and Germany, controversies around the above themes exist in both countries. Moreover, they appear to be played out anew in every new case, sometimes more, sometimes less controversially. England's and Germany's HTA systems thus exhibit an interesting similarity in that controversies around what constitutes good science and good evidence are occurring. According to Majone this is because:

When the issues under discussion require complex patterns of reasoning and large amounts of data of doubtful reliability and relevance, explicit rules of evidence become particularly important (Majone, 1989, p. 10).

The so-called 'rules of evidence' (Majone, 1989) include distinctions between different forms of evidence such as in "[...] the judicial law of evidence with its sophisticated distinctions among proofs of facts, testimony, hearsay, presumptions [...]" (Majone, 1989, p. 10). As shown in the following sections, the content of the controversies in the institutionalised HTA arenas in England and Germany differ from each other in substance, but they have in common that they are controversies about the 'rules of evidence' that Majone (1989) ascribes a great importance to in the policy analysis process.

The empirical examples highlight the similarities and differences of the controversies about science and evidence in England and Germany. The fact that such controversies exist, albeit in different formats, suggests that the conceptualisation and operationalisation of evidence is an important factor that determines the decision on the price (in Germany) and on the availability (in England) of a new pharmaceutical product. The controversies are not unique to HTA systems that are relatively new, nor do they seem to be fully resolved in countries that introduced HTA systems over a decade ago, despite their constant re-developments. This suggests that HTA paradigms are not static, but are confirmed and articulated anew in every decision-making case. Evidence, even the 'same' evidence in the form of RCTs, is not interpreted in the same way in different contexts. How it is interpreted depends on the institutional context in which it is interpreted as well as on the values and principles that are embedded therein.

**TABLE 6.1. - Themes Emerging from the Data (Ranked in order of Prevalence)**

Themes	Properties/questions of the theme in England	Properties/questions of the theme in Germany
1. <i>Permissibility, quality and validity of evidence</i>	<ul style="list-style-type: none"> <li>• What is accepted as evidence?</li> <li>• Evidence applicable to UK clinical practice?</li> <li>• What does the evidence say about the product?</li> <li>• Does it meet the decision-making criteria?</li> </ul>	<ul style="list-style-type: none"> <li>• What is accepted as evidence?</li> <li>• Does the available evidence reflect marketing authorisation and clinical guidelines?</li> <li>• What does the evidence say about the product?</li> <li>• Are the presented clinical endpoints patient relevant?</li> </ul>
2. <i>Choice of comparator product</i>	<ul style="list-style-type: none"> <li>• Is the choice of comparator reflective of UK clinical practice?</li> </ul>	<ul style="list-style-type: none"> <li>• Is the choice of comparator reflective of the current standard alternative?</li> </ul>
3. <i>Patient population subgroup divisions</i>	<ul style="list-style-type: none"> <li>• Do the subgroup divisions adequately reflect the patient groups likely to receive the treatment in routine clinical practice?</li> </ul>	<ul style="list-style-type: none"> <li>• Do the subgroup division adequately reflect the patient population(s) for which the product is licensed?</li> </ul>
4. <i>Operationalisation of criteria for HTA decision-making and role of algorithms</i>	<ul style="list-style-type: none"> <li>• Algorithm is applied: Cost effectiveness threshold in terms of incremental cost effectiveness ratio (ICER) per quality-adjusted life year (QALY)</li> </ul>	<ul style="list-style-type: none"> <li>• What is the algorithm for categorisation of added benefit applied by IQWiG in assessment process?</li> <li>• Algorithm applied by the FJC not transparent (<b>stakeholder opinion</b>).</li> </ul>
5. <i>Suitability of paradigms for 'special' cases such as chronic diseases</i>	<ul style="list-style-type: none"> <li>• How to make decisions in cases where natural progression of the disease is uncertain and patients live with the illness for a long time?</li> </ul>	<ul style="list-style-type: none"> <li>• Questions around the applicability of patient relevant endpoints such as mortality in disease indications where natural progression of the disease is uncertain and patients live with the illness for a long time (<b>stakeholder opinion</b>)</li> </ul>
6. <i>The question of political power and influence: Public pressure and the distribution of bargaining power of stakeholders</i>	<ul style="list-style-type: none"> <li>• What is the effect of public pressure in the form of media and patient campaigns on the final result of HTAs?</li> </ul>	<ul style="list-style-type: none"> <li>• Does the differential distribution of bargaining power of stakeholders impact on the final result of HTAs? (<b>stakeholder opinion</b>)</li> </ul>

The presence of similar controversies about the rules of evidence in Germany and England highlights the pivotal role evidence plays in determining pharmaceutical coverage decisions. The substantive differences in the controversies highlight that this pivotal role should not be equated with the assumption that the same evidence means the same thing to everybody concerned. The ten case studies embedded in this thesis exemplify how the same evidence can be interpreted differently in different contexts, thus giving rise to varying degrees of comparable outcomes in the ultimate decisions.

## **6.2. Theme One: Permissibility, Quality and Validity of Evidence**

One of the reasons why the case of HTAs of pharmaceuticals lends itself particularly well to a comparison of what determines their outcome is that its evidence base is, broadly speaking, comparable. The evidence base in HTA processes is usually the same as that used in the process of acquiring a license under the centralised procedure of the European Medicines Agency (EMA)<sup>26</sup>. This includes randomised controlled trials (RCTs). In contrast to other policy areas such as the environment where regional environmental impact studies ahead of a policy proposal hold a very localised character, the evidence base of new pharmaceutical products is largely the same when it enters national HTA structures. Thus, when examining how decisions on benefit assessments are reached one can control for the variance in evidence and assume that observed differences or similarities are the outcome of other factors. As will be seen, what counts as evidence in the first place and how it is interpreted depends on the dominant national HTA paradigm.

One of the first and most important steps in assessing and appraising a new pharmaceutical product is the critical assessment of the available evidence in order to decide whether it offers answers to the relevant paradigmatic decision questions. This process involves making decisions about which patient populations the medicine is indicated for, which patient population it is likely to be used in, what the current treatment alternatives are and what issues need to be resolved during the appraisal process. The ability of an HTA body to answer the questions it has set itself depends on the evidence that is presented by the pharmaceutical manufacturer. The Evidence Review Group (ERG) at NICE, and IQWiG as a body that is commissioned to carry out

---

<sup>26</sup> See chapter 4 for an outline of how the licensing process works.

assessments by the FJC, look at different features of the evidence in order to decide a) whether the evidence available provides answers to the questions one is interested in, i.e. by being of good quality and comparing the product with the ‘right’ treatment alternative, b) if yes, what the results presented in the evidence suggest about the product and c) how to proceed if the evidence is not powered to provide answers to the relevant questions.

One of the first tasks of an assessment group is thus to decide which evidence is permissible in the first place. This is necessary because there are different types of evidence ranging from clinical trials to observational studies or patient surveys. Scientifically speaking, the type of evidence has an impact on the certainty with which one can assume that the observed outcome under the new treatment is a direct result of the treatment rather than of exogenous factors. In clinical settings RCTs are assumed to be the most valid path for demonstrating the causality of a desired treatment outcome. In designing RCTs one tries to control for as many patient characteristics as possible in order to, ideally, compare two identical patient groups, with one receiving the medical intervention one is interested in and the other receiving a placebo or a comparator product (NICE, 2014b). Other forms of evidence such as patient surveys, observational studies or indirect comparisons give rise to more challenges in that they are open to bias, subjectivity and methodological weaknesses. The paradigmatic decision on which evidence is acceptable to a HTA body is vital in understanding differences and similarities in the outcome of pharmaceutical benefit assessments.

Before exploring the empirical examples of how NICE and the FJC appraise the same evidence, a brief overview of what the methodological guidelines say about the permissibility of evidence is helpful to provide a backdrop against which to understand the analysis. The main points of this overview are presented in table 6.2. In terms of permissibility of evidence, the FJC and NICE have in common that they subscribe to a ‘hierarchy of evidence’ that is frequently found within HTA structures. According to NICE: “Study types [are] organised in order of priority, based on the reliability (or lack of potential bias) of the conclusions that can be drawn from each type” (NICE, 2014b).

For the appraisal of pharmaceuticals this means that both the FJC and NICE have expressed their preference for direct head-to-head clinical trials which take the form of RCTs (NICE, 2008, p. 15; G-BA, 2013). However, the rigour by which the FJC and NICE subscribe to the supremacy of RCTs differs. The NICE Guide to Methods of Technology Appraisal states that:

Non-RCT, both experimental and observational, evidence will be required, not just for those situations in which RCTs are unavailable, but also to supplement information from RCTs when they are available. The problems of confounding, lack of blinding, incomplete follow-up [...] will usually be much worse on non-randomised studies than in RCTs. But in some circumstances, evidence from those studies will be needed in addition to RCT data [...]. In the absence of valid RCT evidence, evidence from studies least open to bias will be considered preferentially with reference to the inherent limitations of the specific design (NICE, 2008, p. 16).

This excerpt of the NICE Guide to Methods of Technology Appraisal signals the preferential use of RCTs as an evidence base whilst not excluding the potential contribution from non-RCT evidence despite the “limitations” (NICE, 2008, p. 16) that such data might bring with it.

In Germany, the FJC divides evidence into seven different levels in its Rules of Procedure, namely:

1. I a systematic reviews of studies of evidence level I b
2. I b randomised controlled studies
3. II a systematic reviews of evidence level II b
4. II b prospective comparative cohort studies
5. III retrospective comparative studies
6. IV Case series and non-comparative studies
7. V Associative observations, [...], descriptive representations, single case reports, expert opinions not underlined by studies, [...], reports of experts committees (G-BA, 2013, p. 94)<sup>27</sup>.

The FJC further specifies: “[...] if it is impossible to carry out studies at the highest level of evidence, then the best available evidence level must be accepted [...]” (G-BA, 2013, p. 94). Adding to this, IQWiG explains that the basis of a benefit assessment is the proof of causality – i.e. the benefit experienced by the patient when taking a new medicine is caused by him/her taking this medicine and not by other confounding factors – which is why, most of the time, RCTs will be considered the only studies that are powered to show this causality (IQWiG, 2008, p. 12). The rigorous division into seven evidence levels with the specification that the lower ranking level of evidence will be considered in the absence of the highest ranking evidence indicates that the hierarchy of evidence in Germany is operationalised much more as a categorical matrix in which the evidence levels are viewed as separate from each other and impermeable

---

<sup>27</sup> Translation provided by the author of this study.

between the levels. In contrast, NICE's operationalisation of the evidence hierarchy indicates that evidence levels are more permeable and viewed as contributing to each other, notwithstanding the general preference for RCTs. Thus, while the emphasis in terms of what constitutes evidence, or at least good-quality evidence, lies on RCTs in both Germany and England, the above remarks offer a first insight into the rigour by which the preference for RCTs is applied and to what extent other forms of evidence are permissible.

Even when HTA systems stipulate a preference for RCTs one has to look closely to understand the extent to which this preference is followed and what other evidence might be permitted. In Germany the adherence to RCTs as the 'gold standard' of pharmaceutical benefit assessment seems to be very strict, with IQWiG's General Methods specifying that the highest level of certainty in a study (i.e. proof of one of the six categories of additional benefit) can only be reached if *two* RCTs of comparable good quality show similar results on clinical data (IQWiG, 2011a). In contrast, a similar requirement in terms of the number of RCTs needed to attain comparable certainty cannot be found in NICE's Guide to Methods of Technology Appraisal. NICE instead labels views by patients and experts as 'evidence' (NICE, 2008, pp. 17-19) which in turn suggests a wider permissible remit in what constitutes evidence.

### **6.2.1. Permissibility of other Types of Evidence**

The preceding analysis gives rise to the question of what happens in cases in which the available evidence is of poor quality or not considered suitable under a given paradigm. In other words, the question arises whether and what types of other evidence are permissible under a given paradigm when the available evidence does not yield answers to the questions posed within the respective decision-making paradigms. The principle of the 'hierarchy of evidence' that is present in the methods guidelines for all three institutions examined here would suggest that evidence of the next highest ranking is sought when evidence of the highest ranking is insufficient or unavailable. However, in practice the hierarchy of evidence is adhered to differently by different HTA institutions, which suggests that contrasting rules of evidence are at play.

Beyond the operationalisation of certain types of evidence such as RCTs, the permissibility of other, lower-ranking, types of evidence plays an important role in determining pharmaceutical coverage decisions. This comes back to the question of

what is regarded as appropriate and good-quality evidence. As outlined above, RCTs are preferred by IQWiG, the FJC and NICE alike. Beyond this, NICE's Guide to Methods of Technology Appraisal (NICE, 2008) and its assessment documents suggest it is more open to recognising additional and alternative sources of evidence such as patient and clinical expert opinions. NICE (2008) goes considerable lengths to explain why patient and clinical expert views can and should be considered as *evidence*. According to NICE,

patient evidence can include views, assessments, evaluations of individual patients, individual carers, groups (such as groups of patients, carers or voluntary organisations that represent patients). [...] Patient evidence refers to any information originating from patients and/or carers that may inform the appraisal of a technology [...]. [...] Patients are a [...] unique source of expert information about the personal impact of a disease and its treatment which can [...] enable the realistic interpretation of clinical and economic data [...]. Patient evidence can identify limitations in the published research literature; in particular, the failure to capture the true concerns of individual patients related to HRQL over and above measurements using standardised instruments [...] (NICE, 2008, p. 22).

The fact that NICE labels the views and statements by patients and other experts as *evidence* offers an indication of the permissibility of these views to inform the decision-making of the Appraisal Committee. Thus, while RCTs are preferred, patients' and clinicians' views are heard and considered on a case-by-case basis within NICE's decision-making paradigm.

A similar regard for patient and clinical expert views as *evidence* cannot be found in IQWiG's assessment documents and methodology papers or the FJC's guidelines. The General Methods of IQWiG highlight that the basis for a benefit assessment is the proof of causality (IQWiG, 2011a, p. 8), i.e. the scientific proof that an improvement or change of a patients' well-being is caused by the medicine that is undergoing assessment. Since RCTs are the most reliable source of demonstrating causality they are considered the highest form of evidence in IQWiG's decision-making paradigm (IQWiG, 2011a). Patients are invited to submit their views in the form of answers to a questionnaire (IQWiG, 2011a), but the assessment documents analysed as part of this thesis do not give insight into how this views are being considered. As we shall see in the next chapter, in Germany a benefit appraisal hinges on the so-called patient relevant endpoints that need to be included in a clinical trial and that need to show significant effects in order for a new pharmaceutical to be categorised as offering

a significant added benefit. The Ordinance for Assessing the Benefit of Pharmaceuticals (Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V, AM-NutzenV) specifies that:

the benefit of a pharmaceutical is the patient-relevant therapeutic effect, in particular in respect of the improvement in the state of health, the reduction of the duration of the disease, longer survival, the reduction of side-effects or an improvement in quality of life (BMJV, 2011).

The requirement for patient relevance is firmly anchored in how IQWiG and the FJC operationalise clinical effectiveness results, namely by assessing only those endpoints that are deemed patient relevant. Even though the phrase ‘patient relevance’ gives rise to connotations of patient preference, it should not be equated with the latter. Indeed, it is one of the more puzzling elements of the German pharmaceutical benefit decision-making paradigm that what is considered relevant by the patients is not always considered as *patient relevant* under the decision-making framework. An interviewee describes patient relevance in the following way:

[...] we would be interested in what the patients notice, what do the patients feel? This means [...] that everything that is connected to morbidity is patient relevant, everything that is connected to side effects [...] is patient relevant [...]. [...] there are surely things that the patients feels but where we would say [...] this is not a medical problem [...] there was the discussion about the insulin analogues and the question whether [...] they are easier to take [...] here I would say [...] the patients feel this but it is nothing that is allowed to lead to costs in the SGB V [...] (Interviewee No. 23, 2013, p. 5).

The benefit assessment in Germany is so closely tied to the requirement for showing patient relevance that it seems to override other potential considerations as well as the possibility of permitting other levels and sources of evidence when RCTs cannot provide answers to the questions sought. The default position of IQWiG and, albeit to a lesser extent, of the FJC seems to be the idea that if a RCT cannot provide statistically significant patient relevant endpoint results, then other forms of evidence such as indirect comparisons, patient and expert views can definitely not produce the missing information or proof of added benefit, hence the benefit will not be proven or it will be considered ‘non-quantifiable’. It seems to result in a situation in which the patient relevance of a clinical effectiveness study is a prerequisite for other forms of evidence to be considered. Thus, when reading the assessment documentation of

IQWiG and the FJC we find ourselves immersed in detailed discussions about whether endpoints are patient relevant or not.

Prominent examples of a discussion of patient relevance are the cases of Telaprevir (IQWiG, 2012c) and Boceprevir (IQWiG, 2011b) in which IQWiG was reluctant to accept the patient relevance of the main clinical endpoint, namely sustained virological response (SVR) because, as IQWiG argued, it is a laboratory endpoint that does not give rise to meaningful conclusions about the effect on individual patients. This position was met with fierce criticism by clinical experts who pointed out that SVR is a worldwide accepted endpoint in hepatitis C and that it was internationally accepted that a SVR equated to healing the patient from his/her infection (G-BA, 2012a).

The principle of patient relevance is worthy of a discussion in its own right in a later chapter of this thesis. For the purposes of this chapter it is important to recognise that patient relevance in Germany is closely connected to the decision-makers' operationalisation of clinical effectiveness, which offers striking differences to other HTA decision-making paradigms. This is a further indication for the hypothesis that what is considered as evidence and what types of evidence are permitted in the first place is closely interwoven with how elements of HTA frameworks such as clinical and cost effectiveness are conceptualised and operationalised in a given health care context. The values, principles and priorities of a given health care system and the society that benefits from its services are embedded in the HTA systems that policymakers construct, thereby giving rise to the different rules of evidence outlined in this chapter.

In the following section I introduce the cases of Cabazitaxel, Eribulin and Ipilimumab as cases in which, on balance, the same evidence in the form of RCTs was accepted by NICE, the FJC and IQWiG. Despite slightly different foci in the appraisal discussions the institutions came to broadly similar conclusions about the evidence that was presented.

**TABLE 6.2. – Permissibility of Evidence and the Hierarchy of Evidence**

What constitutes evidence?	NICE	FJC/IQWiG
1. Hierarchy of evidence	Preference for randomised controlled trials (RCTs)	Preference for randomised controlled trials (RCTs)
2. In the absence of good-quality RCT data/permissibility of other forms of evidence	<ul style="list-style-type: none"> <li>• “[...] evidence from studies least open to bias will be considered [...]” (NICE, 2008, p. 16)</li> <li>• Non-RCT evidence allowed</li> <li>• Other forms of evidence permitted, i.e. patient and expert views also labeled evidence (NICE, 2008, pp. 17-19)</li> </ul>	<ul style="list-style-type: none"> <li>• “[...] best available evidence level must be accepted [...]” (G-BA, 2013, p. 94)</li> <li>• In theory other forms of evidence permitted, in practice this is difficult to implement due to criteria of proof of causality and patient relevance</li> </ul>
3. Hurdles/thresholds	Clinical effectiveness essential but positive recommendation depends on <b>cost effectiveness as expressed in incremental cost effectiveness ratios (ICERs)</b>	Positive benefit assessment hinges on results of <b>patient relevant endpoints</b>

### 6.3. Theme One: The Cases of Cabazitaxel, Eribulin and Ipilimumab

The cases in which NICE, the FJC and IQWiG came to different assessment decisions constitute a large part of this thesis as they provide insights into how the different paradigms are operationalised. However, the cases in which the HTA organisations came to broadly similar conclusions<sup>28</sup> about a product offer equally rewarding opportunities to learn about how paradigms are applied in practice. This became clear during the process of tracing how and why the final decisions on the same pharmaceutical products were made in England and Germany. The process gave rise to the observation that a similar assessment of the weaknesses and strengths of a particular evidence base can result in similar decisions on the benefit of a product even if different

<sup>28</sup> For comparative purposes ‘broadly similar conclusions’ are conceptualised as either positive or negative recommendations. This means that the benefit categories ‘major, significant and marginal’ additional benefit within the German HTA paradigm are operationalised as positive recommendations, whereas the lower three benefit categories are operationalised as negative recommendations. More details on this can be found in chapter 3.

decision-making criteria, or rules of evidence, are applied. This was the case for Cabazitaxel, Eribulin and Ipilimumab. The reason for this seems to lie with the extent of the certainties or uncertainties that a particular piece of evidence gives rise to. That is to say that the quality of evidence and the resulting outcomes (both negative and positive) might be so convincing or uncontroversial that applying different paradigmatic criteria to the evidence will not make a difference in determining the final outcome. This suggests that the good quality or poor quality of a form of evidence can be so extensive that the paradigms that are applied to it do not play a big role.

Before explaining the above in more detail, an overview of the outcome of the dependent variables in the mentioned cases, the salient issues that were raised as well as the main differences and similarities in the appraisal between NICE and the FJC/IQWiG is presented in tables 6.3.-6.5.

**TABLE 6.3. – Case Study: Cabazitaxel (NICE, 2012b; G-BA, 2012c; IQWiG, 2012)**

(Indicated for: Hormone-refractory metastatic prostate cancer previously treated with a docetaxel-containing regimen)

	NICE	FJC
<i>Dependent variable: Outcome of benefit assessment</i>	Not recommended ( <b>Negative outcome</b> )	Different benefit appraisals for two different patient populations ( <b>Negative outcome overall</b> ): <ul style="list-style-type: none"> <li>- Best-supportive care population: Indication for marginal additional benefit</li> <li>- Docetaxel-re-therapy population: No additional benefit substantiated</li> </ul>
<i>Reasoning/discussions/topics raised/public context</i>	<ul style="list-style-type: none"> <li>- Uncertainty about robustness of ICER</li> <li>- ICER too high at £87,500/QALY</li> <li>- Effective, life-extending treatment but too much additional weight would have to be put on QALYs to make it an appropriate use of NHS resources</li> <li>- 1 RCT (TROPIC)</li> <li>- Major concerns around adverse effects</li> <li>- Following consultation on the appraisal consultation document (ACD) the manufacturer provided additional evidence to justify utility values for stable and progressive disease</li> </ul>	<ul style="list-style-type: none"> <li>- Transferability of study results to clinical practice was questioned (patients have more co-morbidities and are older in reality)</li> <li>- Data missing for docetaxel-retherapy population</li> <li>- Additional benefit (in endpoint of overall survival) yes, but have to be weighed against severe side effects</li> <li>- No health-related quality of life (HRQoL) data available = problematic (IQWiG, 2012, p. 48)</li> <li>- Discussion about division of patient population</li> <li>- Operationalisation of categorisation thresholds unclear (G-BA, 2012c, p. 121)</li> </ul>
<i>Similarities in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- Accepted the same RCT (TROPIC)</li> <li>- Overall survival improvement statistically significant</li> <li>- Major concern was adverse effects (NICE, 2012b, p. 35)</li> <li>- Considerable uncertainty about utility values used in cost effectiveness model because HRQoL not available</li> </ul>	<ul style="list-style-type: none"> <li>- Accepted the same RCT (TROPIC)</li> <li>- Overall survival improvement statistically significant</li> <li>- Major concern was adverse effects, these were weighed against endpoint of overall survival which in turn led to a lower benefit category</li> <li>- Considerable weakness that HRQoL data was not available</li> </ul>
<i>Differences in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- Accepted progression free survival as secondary endpoint</li> <li>- Questions of equality between prostate cancer and breast cancer patients were raised by patient groups</li> <li>- Manufacturer appealed the decision: Appeal was rejected</li> </ul>	<ul style="list-style-type: none"> <li>- Did not accept progression free survival as a patient relevant endpoint</li> <li>- Questions of the categorisation of added benefit were raised. Cabazitaxel initially received the same categorisation of added benefit as Abiraterone despite having more adverse effects.</li> </ul>

**TABLE 6.4. – Case Study: Eribulin (NICE, 2012c; G-BA, 2012d; IQWiG, 2012a)**  
 (Indicated for: The treatment of locally advanced or metastatic breast cancer)

	NICE	FJC
<i>Dependent variable: Outcome of benefit assessment</i>	Not recommended ( <b>Negative outcome</b> )	Different benefit appraisals for two different patients populations ( <b>Negative outcome overall</b> ): <ul style="list-style-type: none"> <li>- For patients who cannot be re-treated with regimens containing taxane or anthracycline (different types of chemotherapies): Hint of a marginal additional benefit</li> <li>- For patients who can be re-treated with the other chemotherapies: Additional benefit <b>less than</b> that of the appropriate comparator because of more significant harmful adverse effects</li> </ul>
<i>Reasoning/discussions/to pics raised/public context</i>	<ul style="list-style-type: none"> <li>- ICER of £68,600/QALY likely to be underestimated because of concerns over toxicity profile</li> <li>- Less well tolerated than comparators</li> <li>- No HRQoL collected = important omission</li> <li>- Manufacturer’s model underestimated costs and disutilities of adverse effects</li> <li>- End-of-life criteria not met, overall survival gain less than 3 months</li> <li>- Manufacturer handed in a patient access scheme (PAS) at the beginning of the process</li> <li>- Small number of patients contained in the subgroup analyses of the clinical trial = clinical data in support of Eribulin unconvincing</li> </ul>	<ul style="list-style-type: none"> <li>- Higher risk of severe adverse effects than with the comparator</li> <li>- Overall survival gain: yes, but no HRQoL data, hence lower benefit category</li> <li>- Uncertainty due to the fact that only one clinical study is available and this included a small number of patients</li> </ul>
<i>Similarities in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- Accepted the same RCT (EMRBACE)</li> <li>- Concerns over toxicity profile were high</li> <li>- No HRQoL data = important omission</li> </ul>	<ul style="list-style-type: none"> <li>- Accepted the same RCT (EMBRACE)</li> <li>- Concerns over toxicity profile were high</li> <li>- No HRQoL data = important omission</li> </ul>
<i>Differences in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- Patients acknowledged there were concerns over side effects and the lack of HRQoL</li> <li>- Manufacturer appealed the decision: Appeal was rejected</li> </ul>	<ul style="list-style-type: none"> <li>- Doctors very skeptical about the benefit of Eribulin</li> </ul>

**TABLE 6.5. – Case Study: Ipilimumab (NICE, 2014c; G-BA, 2012e; IQWiG, 2012b)**

Indicated for: Previously treated advanced (unresectable or metastatic) melanoma

	NICE	FJC
<i>Dependent variable: Outcome of benefit assessment</i>	Recommended ( <b>Positive outcome</b> )	Significant additional benefit ( <b>Positive outcome</b> )
<i>Reasoning/discussions/topics raised/public context</i>	<ul style="list-style-type: none"> <li>- ICER £28,600-£47,900/QALY</li> <li>- Initially not recommended</li> <li>- End-of-life criteria met: It's a life-extending end-of-life treatment</li> <li>- Manufacturer submitted a patient access scheme (PAS)</li> <li>- Treatment associated with long-term survival for a small number of patients, but no patient characteristic or biomarkers can be identified to identify these patients prospectively</li> <li>- Appraisal Committee: Innovative therapy because of few advances in the area of melanoma in recent decades; currently the treatment option for these patients is enrolment in clinical trials. Significant innovation for a disease with a high unmet clinical need</li> <li>- Clinicians: this is a step-change, first new treatment in 30 years that may offer clinical benefit and possible long-term survival gain</li> </ul>	<ul style="list-style-type: none"> <li>- Stakeholders opinions in hearings: This is an innovative therapy which should be recognised. Innovative in the sense that its mode of action is new in that it is an immunotherapy that does not directly impact on the tumour but on the immune system's ability to fight it. Special situation in that there is a lack of available alternatives for the patients.</li> </ul>
<i>Similarities in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- Overall survival benefit significant, provides at least an additional 3 months compared with current NHS treatment</li> </ul>	<ul style="list-style-type: none"> <li>- Overall survival benefit is a significant improvement to currently available treatment alternative of 'best-supportive care'</li> </ul>
<i>Differences in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- Focused on innovation as a justification for recommending a treatment whose ICER is higher than that usually considered acceptable within NICE's paradigm</li> </ul>	<ul style="list-style-type: none"> <li>- Innovation not specifically mentioned</li> </ul>

### 6.3.1. Discussion

Tables 6.3.-6.5. show that only Ipilimumab for previously treated advanced (unresectable, i.e. not operable, or metastatic) melanoma received a positive recommendation by both NICE and the FJC. In the cases of Cabazitaxel and Eribulin the appraisals were negative despite FJC's decision to assign a marginal benefit to one of two patient subgroups in both cases. Since the 'marginal benefit' category is the lowest of the positive benefit categories in Germany and the appraisal for the second patient population was very low in both cases (i.e. in Eribulin's case the benefit for the second patient population was less than the appropriate comparator), we observe a situation in which both NICE and the FJC came to, broadly speaking, similar conclusions about the available evidence despite applying different criteria to the evidence.

Cabazitaxel and Eribulin represent new chemotherapies for metastatic prostate cancer and breast cancer respectively. In both cases NICE (2012b; 2012c) and the FJC (2012c; 2012d) accepted the same RCT as the available evidence. The assessment of the weaknesses of the evidence presented in the RCTs was similar in both cases and for both institutions. In both cases there was significant concern around the adverse effects. This, along with the lack of health-related quality of life (HRQoL) data, made a positive appraisal of the two substances challenging. Even though both NICE and IQWiG accepted that the overall survival improved when patients received treatment with Cabazitaxel and Eribulin respectively, the extent of overall survival was not deemed significant enough to justify a positive appraisal because the uncertainties as well as the incidences of adverse effects outweighed the overall survival benefit.

NICE and the FJC justified their decisions on Cabazitaxel and Eribulin in the context of their HTA paradigms. For NICE this meant that although an overall survival benefit was acknowledged, the threshold for a positive recommendation within the cost effectiveness paradigm was not met. In the case of Eribulin the Appraisal Committee at NICE stated that the ICER for Eribulin:

[...] was regarded as a significant underestimate because the concerns about the toxicity profile of eribulin, the uncertainties about health-related quality of life, [...] the use of generic prices to estimate the price of comparators [...] would result in a further increase in the ICER per QALY gained [...] (NICE, 2012c, p. 43).

That is to say that, in addition to concerning toxicity profiles, at ICERs of £68,500/QALY for Eribulin and £87,500/QALY for Cabazitaxel, the increase in overall survival was not enough to justify the recommendation of the two substances.

In the context of the German paradigm of patient relevance, the endpoints of overall survival and adverse effects were accepted as being patient relevant. However, because IQWiG and the FJC operationalise the paradigm by aggregating the results of all endpoints including those of the adverse effects, IQWiG (2012; 2012a) and the FJC (2012c; 2012d) came to the conclusion that the benefit in overall survival was outweighed by the risk of severe adverse effects. In summary, Eribulin and Cabazitaxel represent cases in which NICE and the FJC came to similar conclusions about what the evidence says about the products, but justified the resulting negative decisions by employing the standards of their decision-making frameworks.

Ipilimumab is a cancer treatment, currently indicated for previously treated advanced and non-operable melanoma, that works by activating the immune system to fight cancer cells. In this sense it is different from Cabazitaxel and Eribulin, both of which are chemotherapies that do not work in such a manner. Ipilimumab received a positive appraisal by both NICE and the FJC, mainly due to the fact that the clinical trial showed a statistically significant overall survival benefit for patients on Ipilimumab and the adverse reactions were considered manageable and tolerable. In comparison to the evidence presented in the cases of Cabazitaxel and Eribulin, there were less uncertainties and both HTA bodies were satisfied that the adverse reactions did not outweigh the gain in overall survival. The Appraisal Committee at NICE also concluded “[...] that there was an unmet need for effective therapies in this patients population” (NICE, 2014c, p. 27).

What is interesting about Ipilimumab as a case study is that when Ipilimumab was introduced to the market patients with previously treated advanced melanoma had little to no treatment alternatives. Both in England and in Germany patients were treated according to so-called best-supportive-care regimens, i.e. treatments used in palliative settings in order to alleviate pain. Alternatively, patients had the option of enrolling in clinical trials in the hope of gaining a few months of life expectancy. In the meantime another treatment called Vemurafenib for the group of patients for which Ipilimumab is indicated entered the market. However, at the time at which this analysis was undertaken, there was a distinct lack of alternative treatments. This lack of viable treatment alternatives, along with the severity and low life expectancy of the disease,

arguably contributed to a situation in which the decision-makers in Germany and England were able to recommend Ipilimumab, despite it, for example in England, being over the usually accepted cost effectiveness threshold. Again, this is a different situation from the one observed in the cases of Cabazitaxel and Eribulin in that there are viable treatment alternatives for patients, whereas the patients eligible for treatment with Ipilimumab have no alternative except palliative care.

Ipilimumab also presents an interesting case in that it is the case that received the highest benefit category (benefit category 2 = significant added benefit) in Germany of all ten cases. According to the benefit assessment documentations this was because this categorisation followed the aggregation of both positive and negative results (G-BA, 2012e). In contrast to NICE it is unclear whether other factors such as innovation played a role in the final appraisal decision of the FJC. In its final appraisal determination the Appraisal Committee at NICE:

[...] acknowledged that few advances had been made in the treatment of advanced melanoma in recent years and that ipilimumab could be considered a significant innovation for a disease with a high unmet clinical need (NICE, 2014c, p. 33).

Even though the pharmaceutical manufacturer of Ipilimumab mentioned innovation as one of the positive aspects of the product, it is unclear whether the FJC recognised this as a relevant consideration. According to the manufacturer:

Ipilimumab is a product that does not, like others, affect the tumour directly but it is an immunotherapy that activates the immune systems. [...] In our view this is an innovation (G-BA, 2012e, p. 104).

Considering that the methodological guidelines and interviews with stakeholders in Germany suggest that the aspects that IQWiG and FJC will usually consider are those that are considered patient relevant, i.e. mortality and co-morbidity, it is unlikely that the innovative character of Ipilimumab featured prominently in the FJC's decision-making. However, IQWiG and the FJC were satisfied that the endpoints that were presented in the clinical trials were patient relevant and that the results were, when aggregated, favourable towards a positive recommendation.

Despite the different reasoning that went into the final decisions, Ipilimumab represents a case in which the evidence was, overall, assessed in a similarly positive

fashion by NICE and the FJC. It demonstrates that the cost effectiveness paradigm in England offers flexibility when a given ICER is higher than might usually be considered cost effective. In situations like these decision-makers can refer to additional aspects such as innovation that may ultimately justify the recommendation of the product. However, the German paradigm does not provide for a similar 'contingency plan'.

In summary, the case of Ipilimumab is a special case in that the clinical need for a new treatment was dire and the presented clinical evidence was convincing. In all of the above three cases the decisions came down to a similar assessment of the available evidence. This means that, in these instances, the fact that differing criteria of decision-making were applied, i.e. innovation at NICE in the case of Ipilimumab and an aggregated version of clinical endpoints by IQWiG and the FJC in the case of Cabazitaxel and Eribulin, did not result in significant differences in the final appraisals. The empirical evidence suggests that the outcome in the dependent variable of these cases was comparable because ultimately the decision-makers in England and in Germany felt that the available evidence in the form of RCTs addressed the main criteria and concepts as outlined in the HTA paradigms in an appropriate way. However, the case of Ipilimumab is also an instance in which additional values, i.e. innovation, had to be invoked so that the operationalisation of cost effectiveness matched the significance of the clinical benefits.

The above observations are significant in two ways. Firstly, they suggest that there is a value core within HTA policy paradigms, that is to say certain values and concepts have to be adhered to. This claim arises from the empirical evidence because in Germany the reasoning in the assessment processes for the above cases was centred on the question of patient relevance, whereas in England the reasoning focused on the question of cost effectiveness.

The fact that in Ipilimumab's case NICE based its decision on additional values such as innovation does not contradict the overriding importance of cost effectiveness within the English paradigm. This is because the core of NICE's decision-making reasoning was still cost effectiveness, its default position. That is the focal point from which decisions are taken. In the absence of cost effectiveness thresholds being met NICE has developed a number of additional values that sit just outside the core value of cost effectiveness; they may be brought into the core in order to provide additional grounds of reasoning in certain cases. Thus, and secondly, rather than contradicting the

inference that cost effectiveness is the most important value in NICE's HTA paradigm, the additional, or periphery, values that were invoked by NICE in the case of Ipilimumab suggest a within-paradigm flexibility of reasoning in cases which are not considered cost effectiveness.

As shown in the next chapters a comparable within-paradigm flexibility that is expressed by outlining values that might be invoked if the criteria of the overriding concept are not met is lacking in Germany. If the criterion of patient relevance is not met, Germany's HTA paradigm contains few to no additional values at the outliers of the core of patient relevance that might be invoked. While the policy paradigms of pharmaceutical benefit assessments in Germany and England show similarities in that core and a periphery of values exists, they show differences in the form they take and how decision-making situations are managed when the available evidence does not address the most important decision-making criteria in a satisfactory way.

### **6.3.2. Research Findings**

Cabazitaxel, Eribulin and Ipilimumab illustrate that in cases in which the decision-makers rely on the same forms of evidence and assess the strengths and weaknesses of this evidence in a largely similar way, the result in the outcome of the dependent variable might be similar even if different rules of evidence are applied to the data. The above analysis suggest that this will be the case if a) the same form of evidence is accepted by two HTA bodies and b) the weaknesses of the evidence do not undermine the most important criteria in the value core of the respective HTA paradigm.

For the cases of Cabazitaxel, Eribulin and Ipilimumab, the aforementioned conditions were satisfied. NICE, the FJC and IQWiG accepted the same RCTs as the basis for decision-making. Despite a different focus during the decision-making process, i.e. NICE focusing on innovation and the FJC focusing on the aggregation of patient relevant endpoints in the case of Ipilimumab, the HTA bodies came to largely the same conclusions about the products. The empirical evidence suggests that this was possible because the RCT data addressed the most important decision-making criteria for NICE and the FJC in a convincing way. For Cabazitaxel and Eribulin, NICE found that the incremental costs for the intervention were too high in comparison to the marginal benefits the products offer, even when alternative economic models were

employed. Thus, NICE was able to justify its decision on the grounds of cost effectiveness, or in these cases, the lack thereof.

In the case of Ipilimumab the usual cost effectiveness threshold for recommending interventions was exceeded. However, upon the application of additional criteria such as end-of-life considerations and the question of innovation, NICE was able to recommend Ipilimumab. Whilst NICE's decision-making process in these cases centred on the questions of cost effectiveness, the FJC's primary concern was whether the clinical endpoints of the trial were patient relevant. Being satisfied that they were met, the RCT evidence was deemed appropriate to address the most important questions under Germany's HTA paradigm.

The way in which NICE addressed the issue of cost effectiveness in the above cases demonstrates how the values that are embedded in the wider health care paradigm and the HTA policy paradigm are reflected in the practical application of decision-making criteria. NICE's main focus in the presented cases was the question of cost effectiveness, which has its roots in what is described as 'value for money' in the NHS Constitution (Department of Health, 2013a). NICE's consideration of innovation in the case of Ipilimumab demonstrates that it applies this concept in practice by using it in its reasoning process in cases in which cost effectiveness criteria might not be satisfied. Equally, the fact that the FJC and IQWiG did not accept the question of innovation as a valid consideration in the case of Ipilimumab also reflects how the wider health care and HTA paradigms are put into practice. As outlined in the previous chapter, innovation is not specifically mentioned as a value within the German HTA paradigm. This helps explain why, in the case of Ipilimumab, the FJC and IQWiG did not consider it in its decision-making process.

In conclusion, the cases of Cabazitaxel, Eribulin and Ipilimumab give rise to the following findings:

1. The application of different policy paradigms can lead to similar outcomes if the most important values/criteria of the respective paradigm are safeguarded, thus suggesting that contrasting policy paradigms do not necessarily lead to contrasting outcomes.
2. HTA policy paradigms take different forms in different countries as demonstrated by the fact that contrasting issues were considered by NICE, the FJC and IQWiG when making decisions on Cabazitaxel, Eribulin and Ipilimumab.

3. The issues that were considered by the decision-making bodies in the cases of Cabazitaxel, Eribulin and Ipilimumab reflect the values that are embedded in the paradigmatic constructs of pharmaceutical benefit assessments in Germany and England, which were outlined in the previous chapter.
4. Finally, the above cases illustrate the central importance of the principle of cost effectiveness within the English HTA paradigm and the principle of patient relevance within the German HTA paradigm. The majority of the decision-making process in the two countries focused on these two principles respectively. This suggests that the operationalisation of cost effectiveness and patient relevance plays a central role in shaping the outcome of pharmaceutical benefit assessments. As will be shown in the next chapters, the two concepts can be described as the most important values within the respective HTA paradigms. In paradigmatic terms, they dictate what is considered possible, feasible and desirable within the process of normal decision-making. However, despite the presence of these important values, the case of Ipilimumab also demonstrates that England's paradigm is characterised by a within-paradigm core and periphery of values that allows additional values to be invoked if cost effectiveness thresholds are not met.

Overall, the above cases demonstrate that there are lessons to be learnt from cases in which the outcomes of pharmaceutical benefit assessments are similar in different countries. This is because the in-depth analysis of the decision-making process provides insights into how the decisions were made and which issues were considered. The difference in the issues that were considered can be interpreted as a reflection of different policy paradigms and rules of evidence that are at play. Thus, an important theoretical lesson is that cases in which the outcome of the dependent variable is similar provide as powerful insights when addressing a given research question as cases in which the outcome of the dependent variable is dissimilar.

In the next section I turn to a discussion of the suitability of paradigms for 'special' cases (theme five) as an additional variable that arises from the empirical evidence.

#### **6.4. Theme Five: Suitability of Paradigms for ‘Special’ Cases**

I summarise a theme that was raised in the cases of Retigabine and Fingolimod (see chapter 7) as questions about the ‘suitability’ of the HTA decision-making paradigms for ‘special’ cases such as treatments for chronic diseases. The term ‘special’ is used here for want of a better expression. What is meant is that the cases of Retigabine and Fingolimod gave rise to a set of challenges that are unique to diseases that a) are chronic in kind, i.e. they can span a long timeline and b) still pose unanswered questions in terms of their natural progression and causes. The discussion that follows shows that the questions and arguments that were raised in these cases reflect questions of what evidence is permissible (theme one) for the ‘special’ set of issues and uncertainties that chronic diseases give rise to.

The question of the suitability of the HTA policy paradigms for ‘special’ cases is relevant to the research question because the way in which HTA bodies approach challenges in cases such as chronic diseases might have implications for the likelihood of positive assessments in these circumstances. As the following paragraphs show, the question of the suitability of the HTA decision-making paradigms for these cases also links back to how a HTA body deals with uncertainties in evidence and what kind of evidence is permissible in the first place. It also reflects a more general concern within the academic community about the appropriateness of relying heavily on RCT data to inform health care decision-making (Klein, 2003). This is largely due to the limitations of RCTs to answer questions that go beyond those of clinical effectiveness and safety, for example value judgements that they are not powered to answer.

In the following paragraphs the case of Retigabine is used to illustrate that NICE’s cost effectiveness paradigm seems well equipped to cope with the special challenges that arise in assessing medicines for chronic, long-term, conditions. This seems to be a result of a paradigm that includes a number of ‘contingency’ arguments and values that decision-makers can rely on if a decision cannot be justified on the grounds of cost effectiveness alone. However, the generalisability of the claim that NICE’s cost effectiveness paradigms is better equipped to deal with such situations should be viewed with caution as it is only based on the two cases in which this theme emerged as part of this thesis.

## 6.5. Theme Five: The Case of Retigabine

The case of Retigabine (table 6.6.) serves as an example in which the hard clinical endpoints that are measured in a RCT, for example mortality and morbidity, might not be appropriate to answer questions about the additional therapeutic value that a new medicine offers. NICE specifically recognises the limitations that RCTs might exhibit in that “[...] RCT data are often limited to selected populations and may include comparator treatments and short time spans that do not reflect routine or best NHS practice” (NICE, 2008, p. 15). The empirical research presented in this thesis suggests an additional and equally pressing limitation of the use RCTs, namely the limitations of RCTs to address issues of clinical effectiveness specific to chronic diseases where the scientific knowledge about the natural progression of the diseases is still limited and hence RCTs might not give sufficient insight about the treatment effect of a new medicine.

As outlined in table 6.6. Retigabine is indicated for epilepsy, a chronic disease under which patients suffer from seizures that is caused by “[...] a sudden burst of intense electrical activity” (Epilepsy Action, 2012) in the brain. The three available RCTs for assessing Retigabine’s clinical effectiveness focused on the outcomes “[...] responder rate (proportion of patients experiencing >- 50% reduction in 28-day total partial onset seizure frequency from baseline [...]); proportion of patients achieving seizure reduction categories [...], quality of life scores” (NICE, 2011a, p. 10) and adverse effects. Retigabine “[...] demonstrated statistically significant benefits in terms of responder rate, reduction in seizure frequency and patients achieving freedom of seizures” (NICE, 2011a, p. 9). Despite recognising several limitations to the available cost effectiveness data on Retigabine NICE recommended it as a treatment option while IQWiG and the FJC did not see proof of the added benefit of Retigabine, mainly because the wrong comparator product had been chosen which rendered the available RCTs unemployable for the proof of additional benefit (IQWiG, 2012d; G-BA, 2012f). The reasons for NICE’s and IQWiG’s/FJC’s decision are traced in more detail in the next chapter. In this section the focus lies with taking a closer look at the clinical endpoints that were measured in the RCTs and what clinical experts in Germany and England had to say in relation to the appropriateness thereof.

**TABLE 6.6. – Case Study: Retigabine (NICE, 2011a; G-BA 2012f; IQWiG, 2012d)**

Indicated for: The adjunctive treatment of partial onset seizures in epilepsy

	<b>NICE</b>	<b>FJC</b>
<i>Dependent variable: Outcome of benefit assessment</i>	‘Conditional’ recommendation → only when treatment with 9 other options has failed	Additional benefit not substantiated
<i>Reasoning/discussions/topics raised/public context</i>	<ul style="list-style-type: none"> <li>- ICERs highly uncertain, but recognised a novel mode of action</li> <li>- Provision of new treatment option where others have failed</li> <li>- Compared Retigabine with a number of other alternatives: sometimes it was associated with fewer QALYs and more costs, sometimes with more QALYs and fewer costs</li> </ul>	<ul style="list-style-type: none"> <li>- Problem around deciding what the appropriate comparator was. There are a number of alternatives with no clear rankings, hence the most efficient (cheapest) one (Lamotrigin) was chosen in compliance with the Rules of Ordinance on early benefit assessment</li> <li>- Could not accept any of the presented evidence as none of the trials compared Retigabine with Lamotrigin</li> <li>- In the hearings: discussions about difficulty of designing clinical studies for chronic conditions such as epilepsy</li> </ul>
<i>Similarities in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- Difficult to decide where Retigabine fits in the treatment pathway</li> </ul>	<ul style="list-style-type: none"> <li>- Difficult to decide where Retigabine fits in the treatment pathway</li> </ul>
<i>Differences in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- NICE acknowledged the problem of not knowing exactly where Retigabine would fit in the treatment pathway but saw itself in a position to make a decision on Retigabine because it represented an additional treatment option in situations where other treatments have failed.</li> </ul>	<ul style="list-style-type: none"> <li>- The FJC and IQWiG did not see themselves in a position to make a decision about the additional benefit of Retigabine because there were no studies comparing Retigabine with the appropriate comparator</li> <li>- Pharmaceutical manufacturer took the product off the German market</li> </ul>

The question that arises with regards to the preference for RCTs in the HTA decision-making frameworks is whether these are appropriate to answer the complicated issues of cause and effect in all disease areas regardless of their issue characteristics (Lowi, 1964). Chronic diseases like epilepsy differ from other disease areas in that they pose different challenges due to the longer time that patients live with the disease. Most chronic diseases will develop gradually over time with variations in their severity at different stages of the disease, which may or may not increase the chances for co-morbidity or death. Patients with diseases such as epilepsy or multiple sclerosis will present themselves with very individual and different stages of their diseases, suggesting that a ‘one size fits all’ treatment might be more difficult to prescribe in comparison with other diseases that are acute rather than chronic in nature. Consequentially, there are challenges that arise with regards to what novel treatments need to achieve and how to design RCTs that appropriately measure the treatment effects. This is compounded by the fact that the knowledge about the natural progression of some chronic diseases, including epilepsy, is still limited. One clinical expert who was involved in the appraisal process of Retigabine at NICE phrased this in the following way:

[...] we don’t really know what the drugs do, we don’t really know why people have epilepsy and you could argue that treating epilepsy the way we do is the same as someone who would be treating anaemia blindly [...] you try this, try that, eventually you will get 80%-90% better [...] (Interviewee No. 20, 2013, p. 1).

Arguably, the more uncertainties around the natural progression and occurrence of a chronic disease, the more difficult it will be to design satisfactory clinical trials (Interviewee No. 20, 2013; Interviewee No. 15, 2013) because of the number of variables that remain unknown. Nevertheless, the accepted endpoints that are measured in RCTs on epilepsy medication are responder rates in terms of seizure reduction (Interviewee No. 20, 2013; Interviewee No, 15, 2013). The most prominent outcome is a 50%-seizure reduction rate. However, clinical experts in Germany and in England who were interviewed as part of this thesis outlined concerns they have with these endpoints:

[...] the usual outcome which is used for regulatory authorities is the 50% seizure reduction but the 50% seizure reduction doesn’t really [...] have a lot of

clinical weight [...] it would be like jumping from the 5<sup>th</sup> floor instead of the 10<sup>th</sup> [...] they [the outcomes] have no clinical meaning, they are ok to convince the FDA and the EMA [...] but [...] it won't help me sell the drug to my patients, what I'm looking for is seizure freedom (Interviewee No. 20, 2013, pp. 1-2).

In a similar vein the clinical expert involved in the appraisal process of Retigabine in Germany noted that:

[...] in epilepsy we are talking about [...] a substance that achieved a 50% reduction of seizures, but this is completely irrelevant for the patient. The big qualitative leap would be zero seizures or one seizure, not 5 or 10 seizures. Seizure freedom is the big big endpoint [...] we no longer need the tenth substance that is as good as carbamazepine, we need the eleventh substance that is better than carbamazepine (Interviewee No. 15, 2013, pp. 3-4)<sup>29</sup>.

The above challenges are reflected in NICE's appraisal document on Retigabine in the form of reference to uncertainties and limitations in the data. NICE technology appraisal guidance 232 states that:

[...] The Committee accepted that there may be limitations to the clinical trials data [...] the Committee understood that [...] patient experts valued seizure freedom more than a reduction in the number of seizures. [...] the Committee agreed that the ICERs presented were all highly uncertain because of the limitations in the availability of data [...] (NICE, 2011a, pp. 22-24).

In contrast to NICE, IQWiG and the FJC did not even get as far as recognising the limitations of the clinical data because they dismissed the case outright on the basis that the manufacturer had not compared Retigabine to the appropriate comparator therapy which the FJC had set, namely add-on lamotrigine. The conclusion reached by IQWiG was the following:

[...] in its dossier the pharmaceutical company compared retigabine with lacosamide and thus deviated from the G-BA's specifications. Moreover it provided no adequate justification for this deviation. [...] By choosing a different comparator therapy [...] the pharmaceutical company did not address the question [...]. Accordingly, the studies submitted [...] were not relevant for the benefit assessment – neither for a direct nor for an indirect comparison. Therefore no proof of an added benefit of retigabine in comparison with the appropriate comparator can be inferred [...] (IQWiG, 2012e, p. 1).

---

<sup>29</sup> Translation provided by the author of this project.

Despite uncertainties and limitations NICE still recommended Retigabine on the basis of the evidence it heard from clinical and patient experts and after balancing those with the uncertainties arising from the trials. With reference to the innovation of this technology NICE stated that: “Retigabine has a novel mode of action and therefore could be an important additional treatment option were it to provide response in those people considered resistant to current therapies” (NICE, 2011a, p. 27). This is an example of NICE’s more flexible approach in terms of the evidence base it permits within its decision-making framework. Whereas IQWiG and the FJC dismissed even the available RCTs on the basis that they did not compare Retigabine to the appropriate comparator – in other words they felt that the wrong data had been collected in the first place which rendered the information and the evidence that was crafted from this data useless - NICE acknowledged the weaknesses of the clinical and cost effectiveness data and weighed these up against other evidence and considerations it heard during the appraisal process. How the German HTA institutions addressed the case of Retigabine seems to confirm Majone’s viewpoint that: “Selecting inappropriate data or models [...], can destroy the effectiveness of information used as evidence, regardless of its intrinsic cognitive value” (Majone, 1989, pp. 10-11). In the case of Retigabine the pharmaceutical manufacturer addressed the ‘wrong’ questions in its dossier as IQWiG and the FJC were interested in the effectiveness of Retigabine against a different comparator.

### **6.5.1. Research Findings**

The case of Retigabine serves to illustrate two important points about the presentation of information as evidence in the form of RCTs. Firstly, RCTs’ ability to answer questions appropriately might not always be given, which in turn poses challenges for the decision-making agencies about what information to choose from the available stock of data (Majone, 1989). Secondly, the comparative view on how NICE and the FJC/IQWiG dealt with Retigabine suggests that even evidence of the highest ranking within the hierarchy of evidence might not be considered sufficient evidence if the RCT in question is at odds with other important aspects of the decision-making framework as was the case in Germany. This illustrates that the “ ‘scientific evidence’ ” (Klein, 2003) represented by RCTs might be accepted, declined or re-interpreted in

different HTA settings depending on the wider decision-making paradigms. This affects the ultimate outcome of a pharmaceutical benefit assessment.

Whilst Retigabine was recommended in England, its ‘no proof of added benefit’ categorisation in Germany led to the pharmaceutical manufacturer taking its product off the German market. As a consequence clinicians and patients have to apply to their respective sickness funds to have Retigabine imported from another country if the clinician and the patient feel that this is the right treatment pathway or if the patient is on Retigabine already due to his/her partaking in the clinical trials (Interviewee No. 15, 2013).

In summary, the discussion underlines the following research findings:

1. Rules of evidence matter. The analysis of Retigabine highlights that the rules of evidence, i.e. the way HTA bodies deal with ‘special’ cases, contribute to the outcome of pharmaceutical benefit assessment.
2. HTA paradigms are different in England and Germany. The previous discussion supports this finding as it illustrates how differently HTA bodies approach the challenges in instances of chronic diseases. ‘Special’ cases may challenge the normal decision-making of HTA agencies, but how these agencies deal with such challenges provides valuable insights on the character and flexibility of the paradigms.

In relation to the second research finding, the case of Retigabine suggests that circumstances may arise in which the dominant HTA paradigms may be challenged, challenged in the sense that certain cases of diseases demand a greater flexibility in the articulation and application of the paradigm. By looking at such special cases one can learn about how HTA decision-making paradigms are articulated in order to address case-specific challenges. In the case of Retigabine NICE chose to follow the views of clinicians’ when deciding on where Retigabine would fit in the clinical pathway, whereas IQWiG and the FJC decided they could not assess the drug because the wrong comparator had been chosen. In the case of the latter, even the views brought forward by stakeholders in the assessment hearing process could not change IQWiG’s and the FJC’s decision that it was unable to appraise the product.

## 6.6. Conclusion

In terms of the potential variables that were outlined in table 3.2. the empirical results discussed here suggest that the most important factor that determines the outcome of pharmaceutical benefit assessments is independent variable number three (the instrument settings of HTA, i.e. the rules of evidence). In itself this might not be a surprise as the way data is processed generally has an impact on the outcome. However, it does suggest that the methodological dimension of HTA plays a comparatively more important role than the political and ethical dimensions of HTA in determining their outcome.

Given that one of the goals of HTA is to de-politicise decisions on pharmaceutical coverage, the findings suggest a success when measured against this goal. Except in a couple of exceptional cases, which are discussed in chapter 8, this study does not imply that political factors such as stakeholder influence matter in a significant way, except in instances where the most important principles of the respective paradigms are already upheld in a manner that satisfies the rules of evidence with regards to thresholds and the like. However, while the dominant role of the rules of evidence implies that decisions are, by and large, not subject to political considerations, for policy-makers it also suggests that any concerns that HTAs give rise to in relation to politically and ethically salient issues can only be tackled by ensuring that the rules of evidence reflect one's political and ethical values. This would require re-politicising the issue. As long as the rules of evidence that are articulated in normal decision-making are not at odds with the wider health care paradigm there might not be a need for this. However, in the future policy-makers might be facing the possibility of having to re-politicise a currently de-politicised area.

The next chapter discusses themes two, three and four that emerged from the empirical analysis undertaken in this thesis.

## **Chapter 7**

### **Policy Paradigms in Operation II: The Interpretation of Evidence**

#### **7.0. Introduction**

The empirical evidence analysed as part of this thesis indicates that the decision about which forms of evidence are accepted (theme one) within a given HTA policy paradigm is compounded, if not determined, by more detailed and complex questions pertaining to the available evidence. Seeing as how HTA frequently has a comparative element to it, that is HTA bodies compare a new product's benefits with that of the currently used alternative, the choice of an appropriate comparator product is an important task when carrying out a benefit assessment. In addition to the choice of comparator, the assessment of the clinical effects in sub-populations of the overall patient population for which a pharmaceutical product is licensed forms a substantial part of any benefit assessment in order to determine the relative effectiveness of the product in different patient populations. The consultation documents and stakeholder interviews strongly suggest that HTA bodies address these important questions in different ways (theme two and three). The fact that these evidence questions form the centre of HTAs signals that different 'rules of evidence' are applied when determining the outcome of pharmaceutical benefit assessments. As we shall see in this chapter, the difference between the 'rules of evidence' reflects contrasting HTA policy paradigms in which different values and decision-making criteria are prioritised over others.

This chapter discusses themes two, three and four (table 6.1.) that emerged from the data. Theme two evolves around the role that the choice of comparator plays in determining the outcome of pharmaceutical benefit assessment. The discussion of this theme will highlight that the choice of comparator can impact significantly on whether the available evidence is deemed appropriate in the sense that evidence might be discarded if it is based on the 'wrong' comparator. This has huge implications for how the strength and weaknesses of the evidence is assessed which in turn contributes to the ability of decision-makers to appraise the product in a positive or negative way. The criteria for choosing the comparator product are a reflection of the 'rules of evidence'. Thus, the significance of this issue underscores the research finding that the 'rules of evidence' are the most important factor in determining the outcome of pharmaceutical

benefit assessments. The cases of Fingolimod and Retigabine are employed to illustrate how theme two emerged from the data that was collected.

In the second part of this chapter I analyse theme three of the empirical results, i.e. the role played by divisions of sub-groups of patient populations. This will highlight that NICE, the FJC and IQWiG follow different rules and criteria in deciding which sub-population analyses are relevant and meaningful. However, in contrast to the previous section on the choice of comparator, this section gives rise to the finding that contrasting policy paradigms, or rules of evidence, do not necessarily lead to contrasting outcomes. This will be shown with reference to the case of Telaprevir. Finally, the last section of this chapter is dedicated to a discussion of how HTA thresholds and algorithms are operationalised (theme four) in England and in Germany and how this contributes to the outcome of pharmaceutical benefit assessments.

### **7.1. Theme Two: Choice of Comparator**

Majone refers to data as the “[...] raw materials necessary for the investigation of a problem [...]” (Majone, 1989, p. 46). In pharmaceutical benefit assessments these raw materials can be described as the elements that make up a RCT, i.e. the product under investigation, the comparator product of the control arm of the study and the patient cohort. In order to assess whether these three elements represent the appropriate data to determine a product’s clinical effectiveness the marketing authorisation of a product provides a useful starting point for decision-makers. Notwithstanding the possibility that clinical practice may lead to a situation in which some pharmaceutical products may be used outside their licensed indication, the marketing authorisation for a product provides the legal remit in which a product can and should be used. While the marketing authorisation does not in itself represent a form of data, it impacts on which data is deemed to be appropriate for the purposes of a pharmaceutical benefit assessment. For example, the indication(s) for which a product is licensed will give rise to the current therapy alternatives, i.e. the comparator products currently in use, and the patients for whom the product is indicated will give rise to the patient cohort and its subgroups.

Considering that market authorisation by the European Medicines Agency (EMA) is identical and valid throughout European Member States, one might be led to assume that the data on comparator products and subgroups of patient populations give

rise to similar benefit assessment outcomes in countries with HTA procedures. However, this is not the case. The empirical research conducted as part of this thesis shows that the ultimate outcome of a pharmaceutical benefit assessment will, amongst other factors, be determined by the strictness to which the marketing authorisation is adhered to when making a decision about the comparator product and patient subgroup divisions. A discussion of the cases of Fingolimod, Retigabine and Telaprevir highlights the challenges that arise around the issue of choosing the appropriate data, i.e. the ‘right’ comparator product and patient subgroups, and how different approaches in dealing with those challenges can lead to contrasting appraisals of the same data.

## **7.2. Theme Two: The Case of Fingolimod**

Fingolimod is a treatment that is licensed for multiple sclerosis (MS). MS “[...] is a disease of the nerves in which inflammation destroys the protective sheath surrounding the nerve cells” (EMA, 2011). It is a chronic disease that comes in different forms and levels of severity. One of the most common forms is that of relapsing-remitting MS in which patients go through periods of remissions with no symptoms and periods of relapses in which they suffer from attacks and MS symptoms. Table 7.1. provides an overview of NICE’s and the FJC’s decisions and reasoning in the case of Fingolimod.

**TABLE 7.1. – Case Study: Fingolimod (NICE, 2012d; G-BA, 2012g; IQWiG, 2012e)**  
 Indicated for: Highly active relapsing-remitting multiple sclerosis (MS)

	NICE	FJC
<i>Dependent variable: Outcome of benefit assessment</i>	<p>‘Conditional’ recommendation</p> <p>Recommended as an option for the treatment of highly active relapsing-remitting multiple sclerosis (MS) in adults, only if: they have an unchanged or increased relapse rate or ongoing severe relapses compared with the previous year despite treatment with beta interferon AND the manufacturer provides Fingolimod with the discount agreed as part of the patient access scheme (PAS)</p>	<p>Different benefit appraisal for 3 different patient populations:</p> <ol style="list-style-type: none"> <li>1. Patients with highly active relapsing-remitting MS who are heavily pre-treated: Additional benefit not substantiated</li> <li>2. Patients with highly active relapsing-remitting MS who are not heavily pre-treated: Additional benefit not substantiated</li> <li>3. Patients with rapidly evolving severe relapsing-remitting MS: Clue of a marginal benefit</li> </ol>
<i>Reasoning/discussion s/topics raised/public context</i>	<ul style="list-style-type: none"> <li>- Accepted 2 RCTs (FREEDOMS and TRANSFORMS)</li> <li>- Initially did not recommend it</li> <li>- Appraisal committee made an exceptional case: valuable new therapy, oral formulation is an innovation</li> <li>- ICER £25,000-£35,000/QALY</li> <li>- Manufacturer acknowledged that trial populations did not meet criteria described in marketing authorisation</li> <li>- ERG very critical of the presented evidence, i.e. populations in the clinical trials were broader than those defined in the marketing authorisation. However, accepted that manufacturer’s post-hoc subgroup analyses provided a reasonable approximation to the populations in the marketing authorisation</li> </ul>	<ul style="list-style-type: none"> <li>- Limited the timeline for which the appraisal decision is valid, meaning it will be reassessed once more data is available</li> <li>- Accepted only 1 RCT (TRANSFORMS), FREEDOMS not acceptable because it was placebo-controlled and did not allow for a comparison to a relevant comparator</li> <li>- Division of population according to marketing authorisation means higher uncertainty</li> <li>- Manufacturer highlighted the significance of the oral formulation of Fingolimod as an added benefit</li> </ul>

<i>Similarities in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- Acknowledged that the mismatch between the patient populations in the clinical trials and the marketing authorisation presented uncertainties in the data</li> </ul>	<ul style="list-style-type: none"> <li>- Acknowledged that the mismatch between the patient populations in the clinical trials and the marketing authorisation presented uncertainties in the data</li> </ul>
<i>Differences in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- NICE was not able to make a recommendation for patients with rapidly evolving severe relapsing-remitting MS because analysis that compared Fingolimod with the appropriate comparator was not submitted by the manufacturer</li> <li>- Oral formulation accepted as a benefit of value</li> <li>- Patient experts placed emphasis on loss of independence and implications for employment under MS. They also placed emphasis on the benefit of Fingolimod as an oral medicine as opposed to one that has to be injected.</li> <li>- Clinical specialists said it would provide the most benefit for patients with rapidly evolving severe relapsing-remitting MS because they currently have few treatment options</li> </ul>	<ul style="list-style-type: none"> <li>- FJC was able to assign a positive (marginal) benefit to the patient group with rapidly evolving severe relapsing-remitting MS</li> <li>- Oral formulation not considered because it cannot be expressed as a patient relevant endpoint.</li> </ul>

### 7.2.1. Discussion

Whilst the challenges presented by the RCT data in the case of Fingolimod were described in similar ways by NICE, IQWiG and the FJC, the way in which these challenges were approached by the institutions as well as the ultimate decisions they arrived at are dissimilar to each other. The challenges related primarily to a lack of congruence between the patient population included in the trials and that covered by the EMA license. This in turn gave rise to discussions about the appropriate comparator product for assessing Fingolimod's clinical effectiveness at NICE as well as IQWiG and the FJC. Whilst NICE and the FJC ultimately decided to recommend Fingolimod for use, the recommendation differed in that it was for two different patient subgroups. Whereas NICE recommended Fingolimod,

[...] as an option for the treatment of highly active relapsing-remitting multiple sclerosis in adults, only if:  
they have an unchanged or increased relapse rate or ongoing severe relapses compared with the previous year despite treatment with beta interferon [...]  
(NICE, 2012d, p. 3),

the FJC did not come to the conclusion that there was an added benefit for this patient population, but instead assigned a hint of a marginal added benefit to the patient subgroup of rapidly evolving severe relapsing-remitting MS to Fingolimod (G-BA, 2012g). By contrast, NICE concluded that:

[...] a specific recommendation for the use of fingolimod in this population could not be made because the manufacturer had not submitted an analysis of fingolimod compared with natalizumab in this population (NICE, 2012d, p. 34).

In the case of Fingolimod the two appraisal institutions felt that, based on the available evidence, they could only make a decision that is limited to one of the patient populations covered by the market authorisation. However, the patient population for which NICE and the FJC issued a positive decision is not identical, thereby giving rise to questions about what determined the ultimate decision. It strongly suggests that the way HTA institutions approach challenges of uncertainty and trial design affects the final outcome of a pharmaceutical benefit assessment.

In the case of Fingolimod, NICE, IQWiG and the FJC highlighted that the patient populations in the trials were more broadly defined than in the marketing authorisation (NICE, 2012d; G-B, 2012g; IQWiG, 2012e). This means that more patients with a wider variety of baseline characteristics and previous treatments were included in the RCTs than covered by the market authorisation. This made the extrapolation of statistically significant results difficult. At NICE “the Committee noted that only part of the population covered by the marketing authorisation for fingolimod was considered in the manufacturer’s submission” (NICE, 2012d, p. 23) and “it also heard that the TRANSFORMS and FREEDOMS trials were not powered to assess the efficacy of fingolimod in the subgroups defined by the marketing authorisation” (NICE, 2012d, p. 25). In a similar vein the FJC concluded that the extrapolation of data results for subgroups that fulfill the marketing authorisation criteria is not possible in this case (G-BA, 2012g). The lack of congruence between the patient subgroups as outlined in the marketing authorisation and those patients included in the clinical trials thus presented NICE and IQWiG with similar challenges when it comes to assessing the appropriateness of the data, the raw material, that was presented as part of the benefit assessment process. The way NICE and IQWiG as well as the FJC dealt with these challenges gives important insights into the HTA decision-making paradigms in Germany and England.

Whilst recognising the limitations in the data presented by the manufacturer the ultimate decision of NICE suggests that it was satisfied with the quality of the additional evidence derived from mixed treatment comparison and indirect evidence that was presented in the process. Even though NICE saw limitations with regards to the comparator chosen for “[...] the part of the population covered by the marketing authorisation [...], that is, people with highly active relapsing-remitting multiple sclerosis” (NICE, 2012d, p. 15), it felt that it could make a decision about the clinical effectiveness of Fingolimod for this patient population. NICE criticised that the manufacturer only chose one beta interferon (Avonex) as a comparator product, but ultimately “the Committee concluded that the available evidence shows that people who are treated with fingolimod have lower relapse rates than people treated with Avonex or placebo” (NICE, 2012d, p. 36). However, for people with rapidly evolving severe relapsing-remitting MS NICE did not make a recommendation due to the fact that, as previously mentioned, “[...] the manufacturer had not submitted an analysis of fingolimod compared with natalizumab [the appropriate comparator] in this population”

(NICE, 2012d, p. 34). Despite the challenges presented by the lack of congruence between the patient populations included in the clinical trials and that covered by the marketing authorisation NICE decided to recommend Fingolimod for one of the patient populations on the basis of all available evidence, including oral and written evidence that it heard from patients and clinical experts about the advantages of Fingolimod and its oral formulation. In the end,

the Committee made an exceptional case and recommended fingolimod for the treatment of highly active relapsing-remitting multiple sclerosis in adults [...]. The Committee accepted that fingolimod is a valuable new therapy and that its oral formulation represents innovation in the treatment of multiple sclerosis. The Committee recognised that including all of the benefits of fingolimod suggested by the manufacturer and patient experts in the manufacturer's model could decrease the ICER to a level that would be considered a cost-effective use of NHS resources (NICE, 2012d, p. 32).

The fact that the Committee made an “exceptional case” (NICE, 2012d, p. 32) suggests that the decision was not easy and that there were factors that spoke against the recommendation of Fingolimod. In fact, NICE's draft recommendation in the Appraisal Consultation Document (ACD) was negative due to the uncertainties in the clinical evidence and the resulting unfavourable cost effectiveness estimates (NICE, 2011b). The reversal of NICE's initial recommendation as well as its decision to grant Fingolimod the status of an innovative pharmaceutical product due to its oral formulation are special features of this case. It is important to acknowledge that NICE – on balance of all available evidence – made a positive recommendation despite uncertainties that arose from the divergence between the marketing authorisation of Fingolimod and the design of the clinical trials. Coming back to the ‘hierarchy of evidence’ that was explained in chapter 6, NICE followed the hierarchy in the sense that it extrapolated the information it saw fit from the available RCT data and then looked to other forms of evidence to fill the information gaps and build a stronger evidence base.

In contrast to NICE, IQWiG and the FJC did not accept the results of the indirect comparisons with the appropriate comparator (glatiramer acetate) or any other evidence for that matter. The indirect comparisons presented by the manufacturer were not accepted by IQWiG because the studies did not include the relevant patient cohort, i.e. patients previously treated with a beta interferon (IQWiG, 2012e). This gave rise to contrasting views between IQWiG and the pharmaceutical manufacturer on what

constitutes the ‘best available evidence’ that has to be presented when the highest ranked evidence is not available. The manufacturer saw the indirect comparisons it presented as the best available evidence (G-BA, 2012g, p. 61), whereas IQWiG concluded that this evidence did not answer the decision problem and could therefore not be accepted. According to IQWiG,

[...] every evidence – that is the best available as well as the best possible evidence for a decision problem – must [...] be suitable for answering a given decision problem. If this is not the case then even the [...] best available evidence for a decision problem is irrelevant. This is the case for the indirect comparisons presented by the pharmaceutical manufacturer [...]. These considerations would also apply [...] to direct comparative studies (IQWiG, 2012e, p. 51)<sup>30</sup>.

Even though the FJC Ordinance stipulates that the “best available evidence level” must be accepted in the absence of the highest level of evidence (G-BA, 2013), IQWiG seems to draw a further distinction between the best *available* and the best *possible* evidence. Thus, in this case IQWiG’s and the FJC’s decisions were a result of a strict adherence to the marketing authorisation, which was seen as binding. However, as will be further illustrated by the case of Telaprevir in a later section of this chapter, the use of the marketing authorisation to determine the exact patient populations and current therapy alternatives in a given indication for the purposes of pharmaceutical assessments is not without its challenges. The challenges lie with issues of clinical practice and practicability as well as the divergent goals of the licensing and the appraisal processes. In connection to this, an interviewee of a decision-making body in Germany pointed out:

A [...] topic [...] is the role of the marketing authorisation which is at times a big problem [...] because we are clearly [...] bound by the marketing authorisation in our SGB V [Social Code Book V] and this is difficult for the manufacturers and for us [...] (Interviewee No. 23, 2013, p. 1).

---

<sup>30</sup> Translation provided by the author of this thesis.

### 7.2.2. Research Findings

The case of Fingolimod illustrates the impact that the application of the marketing authorisation has on the appraisal process in that it gives rise to the patient population, its subgroups as well as the comparator products in clinical practice. The case also illustrates how additional evidence of a lesser hierarchical ranking than RCTs may or may not be used by decision-makers to answer a given decision problem. The ability (by law) and the willingness of decision-making institutions to employ alternative sources of evidence in the absence of direct comparisons can have a large effect on the ultimate outcome of a decision. It is for this reason that Fingolimod is recommended for a large part of the patient population covered by the marketing authorisation in England (albeit with conditions around previous treatment and relapse rates) while in Germany it is only categorised as offering a small added benefit for a small part of the patient population. This shows that, when it comes to the role played by the marketing authorisation and the issues it raises, divergent ‘rules of evidence’ (Majone, 1989) can have a very real impact on what determines the outcome of pharmaceutical benefit assessments.

In the case of Fingolimod, the reasoning of the decision-making bodies in Germany and England shows that the approaches to dealing with weaknesses in the evidence base and decisions about the ‘right’ comparator are very different. The process-tracing exercise in this case gave rise to the following findings:

1. It does not just matter what forms of evidence are permissible (theme one) under a given HTA paradigms. How the evidence is interpreted in relation to questions such as the appropriate comparator product impacts on the outcome of the benefit assessments. This is because the decision that the evidence does not include the ‘right’ comparator effectively rules out making an appraisal decision on the basis thereof.
2. The question of the appropriateness of the comparator and the sub-divisions of patient populations are connected to what is considered ‘relevant’ in a specific HTA system. This underlines Cartwright’s and Hardie’s (2012) assertion that evidence has to be relevant to a specific context. Under the English HTA paradigm the question of relevance is operationalised by asking whether the clinical trial data reflects UK clinical practice, whereas under the German

paradigm it is operationalised by asking whether it reflects the marketing authorisation of the given pharmaceutical product.

3. Decision-making criteria are established as part of a paradigm during normal decision-making processes. In these processes HTA bodies find a way of dealing with situations in which the ‘wrong’ comparator was chosen. Here, the case of Fingolimod suggests that NICE is willing to consider evidence other than RCTs such as indirect comparisons and expert submissions in order to make a decision. The study data does not include any evidence that the FJC is willing to consider evidence from sources other than RCTs.
4. Rules of evidence matter. The fact that NICE, the FJC and IQWiG asked similar questions about the evidence base and dedicated a substantial amount of time within the decision-making processes to these questions demonstrates a similar importance of questions regarding the ‘rules of evidence’ in both HTA paradigms. This also shows how paradigms are further articulated and substantiated during ‘normal’ decision-making processes.
5. Different rules of evidence support the theoretical premise that contrasting HTA policy paradigms are at play. The fact that the aforementioned HTA bodies ultimately found different answers to the pressing questions, i.e. with NICE accepting indirect comparisons and IQWiG and the FJC not doing so, indicates that different rules of evidence are applied in pharmaceutical benefit assessment processes in Germany and England.

### **7.3. Theme Two: The Case of Retigabine**

Another example that highlights the importance of the choice of comparator in determining the final outcome in pharmaceutical benefit assessment is the case of Retigabine (see table 6.6.). Even though the EMA has restricted its use to last-line therapy in patients where other treatment options have failed due to the incidence of an adverse reaction that causes abnormal colouring of the skin (EMA, 2013), Retigabine exemplifies how different HTA paradigms affect the way in which similar challenges in the available evidence are dealt with by HTA institutions. In Retigabine’s case this ultimately led to a positive recommendation by NICE (2011a) whereas the FJC (G-BA, 2012f) did not consider itself in a position to make a decision on the additional benefit of this product.

### 7.3.1. Discussion

The main challenge in the case of Retigabine was the judgement around its place in the treatment pathway and the most appropriate comparator. While clinical specialists in Germany and in England agreed that the most likely place of Retigabine in clinical practice would be as a second-line therapy in patients where other treatment options have failed or have been exhausted, the FJC and IQWiG referred to the marketing authorisation of the medicine under which it was also licensed as a first-line treatment (G-BA, 2012f; IQWiG, 2012d). Retigabine's license as a first-line treatment was the reason why the FJC chose Lamortigine, also indicated as a first-line treatment in epilepsy, as the most appropriate comparator. This choice caused much controversy during the HTA process on Retigabine in Germany. This is because there are a number of medicines licensed as a first-line treatment for epilepsy and stakeholders criticised that the FJC chose Lamortigine out of all of these treatments (G-BA, 2012f). Effectively the FJC's decision led to a situation in which it could not make a decision about the additional benefit of Retigabine because none of the available trials and evidence compared Retigabine with Lamortigine. The FJC argued that in the absence of a clear ranking in clinical guidelines of the order in which the various treatment alternatives should be used it was under a statutory obligation to choose the most efficient, i.e. the cheapest, available alternative which was Lamortigine (G-BA, 2012f).

In contrast to the FJC, NICE accepted what clinical experts said about the likely place of Retigabine in the clinical treatment pathway. In light of the fact that epilepsy treatment is very individualised because every patient reacts to treatment differently, NICE accepted a range of alternatives as appropriate comparators and decided to recommend Retigabine once these alternatives have failed to induce a beneficial outcome (NICE, 2011a). One of the main reasons for NICE's decision was the acknowledgement that an additional treatment option would be good for patients who have exhausted all other treatment options (NICE, 2011a). In Germany the existing HTA paradigm prevented the FJC decision-makers from taking a similarly flexible approach to the challenges of assessing Retigabine. One stakeholder summarised the FJC's approach during the early benefit assessment in the following way:

[...] it [the FJC's approach] is very formal [...] it sticks to law and order very strictly and there is little room for practical solutions [...] especially in the case

of Retigabine one sees that NICE has a different viewpoint, they said ok, after those [other treatments] you are allowed to treat with Retigabine, there is a benefit here and we will reimburse it then. We tried to argue the same thing but the answer was, yes, but unfortunately this is not possible in our Social Code Book-driven system [...] (Interviewee No. 15, 2013, p. 2).

The last sentence of the above quote gives a deep insight into the strictness by which the HTA paradigm is adhered to in Germany and how this restricts a certain level of flexibility in some cases. The phrase “[...] this is not possible in our Social Code Book-driven system [...]” (Interviewee No. 15, 2013, p. 2) can be taken as a code for ‘this is not possible under the current HTA decision-making paradigm’. This paradigm clearly states that the marketing authorisation provides the guidance for making decisions about a medicine’s place in the treatment pathway and its relevant comparators. It also clearly stipulates that in the case of more than one suitable comparator the most efficient one has to be chosen (BMJV, 2013).

The FJC’s inability to make a decision on the added benefit of Retigabine due to the lack of available evidence on what it considered the appropriate decision problem is a reflection of the ideas that form the basis of the HTA paradigm. It represents an example of how similar challenges are dealt with very differently depending on the paradigmatic decision-making criteria that apply. In the case of Retigabine this had a large impact in that the availability of the product is different in Germany and England. Following the FJC’s negative decision, the pharmaceutical manufacturer in Germany decided to withdraw Retigabine from the German market. This means that patients who want to receive it and clinicians who want to prescribe it have to apply to individual sickness funds in order to have it imported from another European country (Interviewee No. 15, 2013). This illustrates how paradigms, interpretive frameworks within a given policy area, can have a very real political and social impact.

### 7.3.2. Research Findings

The English HTA paradigm offers comparatively more room for the consideration of evidence other than RCTs which is why, in the case of Retigabine, NICE was able to consider what clinicians had to say about the administration of Retigabine in routine clinical practice.

Faced with the same challenges when presented with the evidence, Fingolimod and Retigabine represent case scenarios in which the HTA bodies dealt with the challenges in a diametrically opposing manner, which ultimately led to different outcomes in the respective benefit assessment decisions. The contrast in the approaches taken in the appraisal of Retigabine is especially striking; whereas the FJC (G-BA, 2012f) and IQWiG (2012d) accepted *none* of the comparators that were included in the available evidence, NICE (2011a) accepted the comparators *and* included a range of comparators that had not even been a part of the direct evidence presented by the manufacturer.

Despite questions about the place of Retigabine in routine clinical practice, NICE was able to recommend Retigabine because it followed what clinicians were saying about the use of Retigabine as an option when other treatments had failed. IQWiG and the FJC on the other hand did not get past Retigabine's license as a first-line therapy. NICE's decision was possible because its paradigm allows for clinicians' viewpoints to be counted as evidence. Even though the German HTA paradigm includes the stipulation that the next best level of evidence has to be used in the absence of the highest level of evidence (see chapter 6), this stipulation does not appear to extend to recognising clinicians' views as acceptable evidence. The example of Fingolimod and Retigabine thus underlines the importance of examining how decision-making criteria such as hierarchies of evidence are operationalised when put into normal practice.

In summary, the discussion on the role of the choice of comparator in determining the outcome of pharmaceutical benefit assessment gives rise to the following research findings:

1. The 'rules of evidence', as expressed through different paradigmatic prisms, matter. The discussion of the cases of Fingolimod and Retigabine highlights that the outcome of pharmaceutical benefit assessments is not solely determined by the evidence input the HTA body accepts, but also by the means (the rules of evidence) by which this input is analysed.

2. The discussion on Retigabine and Fingolimod showed how variables combine in pharmaceutical benefit assessment to produce certain outcomes. In the above cases the application of different rules of evidence ultimately led to different results despite a similar assessment of the weaknesses in the evidence base. This suggests that it is not the evidence base that determines the outcome of pharmaceutical benefit assessment, but the decision-making criteria that are applied to it. By applying different criteria the evidence gains relevance within a given HTA paradigm.

In the next section I elaborate on the third theme that was raised in the majority of the cases. Along with the choice of comparator product and the paradigmatic decision about which evidence to accept the question of the appropriate sub-division of the patient population also plays a role in determining the final outcome in pharmaceutical benefit assessments.

#### **7.4. Theme Three: The Sub-Divisions of Patient Populations**

An important element in designing and carrying out RCTs is the definition of subgroups of patients that may experience a different treatment effect under the medicine that is being tried. The subgroups will usually undergo what is called a subgroup analysis, that is: “A technique consisting of analysing the data from a subgroup separately from those from the overall population studied” (HTA Glossary, 2014a). Subgroup analyses potentially offer decision-makers a closer look at the effects of the treatment which might, for example, lead to restricting the use of the product in question to the patient groups that are most likely to benefit from its use. However, subgroup analyses are not without challenges as dividing the patient population into groups that are smaller than the entire cohort will also affect the statistical significance results that are likely to be obtained in these groups. Despite these challenges that have to be given due consideration, subgroup analyses form an important part of the work HTA institutions perform in assessing and appraising a product’s treatment effects. The extent to which HTA institutions engage in subgroup analyses offers an interesting insight into the paradigmatic differences of how data is used and analysed in different contexts. The case of Telaprevir serves as an example of the important role that the sub-

divisions of patient populations play and how this might impact on the decisions made in relation to assessing the benefit of a new medicine.

### **7.5. Theme Three: The Case of Telaprevir**

Telaprevir is:

[...] indicated for the treatment of genotype 1 chronic hepatitis C in adult patients with compensated liver disease (including cirrhosis):

- Who are treatment naïve;
- Who have previously been treated with interferon alfa [...] alone or in combination with ribavirin, including relapsers, partial responders and null responders (EMA, 2011a).

Table 7.2. provides an overview of the most important points that can be traced by examining the benefit assessment process at NICE, IQWiG and the FJC.

**TABLE 7.2. - Case Study: Telaprevir (NICE, 2012e; G-BA, 2012a; IQWiG, 2012c)**

Indicated for: The treatment of genotype 1 chronic hepatitis

	<b>NICE</b>	<b>FJC</b>
<i>Dependent variable: Outcome of benefit assessment</i>	Recommended	Additional benefit, but not quantifiable
<i>Reasoning/discussion/topics raised/public context</i>	<ul style="list-style-type: none"> <li>- ICERs low at £13,553/QALY for treatment-naïve patients and £8,688/QALY for treatment-experienced patients</li> <li>- Significant features of the treatment: by increasing sustained virological response (SVR) rates Telaprevir increases the chances of stopping the progression of the diseases to more disabling stages of liver disease</li> <li>- Offers treatment opportunities for patients with a type of hepatitis C virus that is most resistant to current treatment alternatives</li> <li>- Public health benefit highlighted: lower transmission rates if more patients achieve SVR</li> <li>- More clinically effective than the current alternatives</li> <li>- Innovation highlighted</li> </ul>	<ul style="list-style-type: none"> <li>- Questions around the sub-divisions of the patient population by IQWiG → the FJC did not follow IQWiG's sub-division into 5 different sub-groups</li> <li>- Questions and controversy around whether the clinical endpoint SVR was patient relevant</li> <li>- IQWiG interpreted SVR as a surrogate parameter for liver carcinoma, not as an endpoint in its own right</li> <li>- The FJC accepted SVR as a patient relevant endpoint</li> <li>- Clinical experts highlighted public health consideration of lower transmission rates</li> <li>- Discussion about operationalisation of benefit categories</li> <li>- Stakeholders argued Telaprevir should receive the highest benefit category because SVR is equivalent to a cure</li> </ul>
<i>Similarities in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- Effective treatment: yes</li> <li>- Stakeholders highlighted effectiveness and innovation of treatment</li> </ul>	<ul style="list-style-type: none"> <li>- Effective treatment: yes</li> <li>- Stakeholder highlighted effectiveness and innovation of treatment</li> </ul>
<i>Differences in assessment &amp; appraisal</i>	<ul style="list-style-type: none"> <li>- No doubt about appropriateness of clinical endpoint SVR</li> <li>- Consideration of public health impact: yes</li> <li>- Accepted that SVR was equivalent to a cure</li> </ul>	<ul style="list-style-type: none"> <li>- Much doubt and discussion about appropriateness of clinical endpoint SVR</li> <li>- Consideration of public health impact: No</li> <li>- Did not accept that SVR is equivalent to a cure</li> </ul>

### 7.5.1. Discussion

While both NICE and the FJC came to an overall positive appraisal of Telaprevir, the reasons for this as well as the discussions that preceded the ultimate decision differed substantially from one another. The differences in the way in which subgroup effects were discussed and analysed within the German and the English context are especially noteworthy as they offer an empirical example of how different subgroup divisions and analyses have a potentially large impact on the ultimate outcome in pharmaceutical coverage decisions.

In its ultimate decision the FJC concluded that there was an added benefit of Telaprevir for treatment naïve and previously treated patients (G-BA, 2012a). However, the FJC also concluded that the extent of the added effect was not quantifiable because HIV and hepatitis B (HBV) co-infected patients and patients with liver cirrhosis were potentially included in the groups for whom an added benefit was accepted (G-BA, 2012a). According to the FJC there was no sufficient data for these subgroups of co-infected patients (as they were excluded from the clinical trials) and patients with cirrhosis which is why no statements of certainty could be made about the primary endpoint of sustained virological response (SVR) (G-BA, 2012a, p. 4). According to the FJC the “scientific data base” (G-BA, 2012a, p. 4) does not permit a quantification of the extent of the added benefit. Thus, the FJC’s caveated decision on Telaprevir was guided by an uncertainty around the treatment effects within certain patient subgroups and whether this skewed the effects in the entire patient population.

It is important to note that in the above conclusion in the case of Telaprevir the FJC did not follow the divisions of the patient population that had been suggested by IQWiG in its assessment of Telaprevir. IQWiG divided the patient population into five sub-populations that, according to IQWiG, more accurately reflected the patients for which Telaprevir is indicated as per its license. According to IQWiG, the appropriate sub-populations and the resulting conclusion about their added benefit were:

- Treatment-naïve patients without cirrhosis (proof of added benefit);
- Treatment-naïve patients with cirrhosis (no proof of added benefit because of lack of useable data);
- Non-responders with or without cirrhosis (hints of an added benefit);
- Relapse-patients with cirrhosis (data uncertain);

- And relapse-patients without cirrhosis (no proof of added benefit) (IQWiG, 2012c, pp. 3-11).

In addition to these subgroup analyses, IQWiG conducted a subgroup analysis on the impact of the virological load of patients at baseline for the sub-population of therapy-naïve patients without cirrhosis. This led to IQWiG concluding that the added benefit of Telaprevir for treatment-naïve patients without cirrhosis was only proven for patients with a high virological load at baseline (IQWiG, 2012c). IQWiG's reasoning for these sub-divisions, while arguably correct on the possible sub-populations that arise out of the license for Telaprevir, was met with considerable criticism by stakeholders in Germany.

The criticism brought up most frequently during the the FJC's hearing procedures on Telaprevir was that the subgroup divisions by IQWiG led to an under-appreciation of the positive treatment effects of Telaprevir (G-BA, 2012a, p. 36). Underlying this criticism is the previous mentioned challenge of subgroup analyses, especially if they are conducted post-hoc, of showing statistically significant results in a much smaller patient population than the full cohort. Moreover, stakeholders raised the issue of the practicability of IQWiG's subgroup divisions:

Practicability plays a big role [...]. Telaprevir is a very good example, [...] IQWiG opened up seven subgroups in total, which were in part even further divided [...] this was too complex for the FJC, I think, and one decided [...] in order to have a comprehensible decision for the clinician [...] to limit the subgroups to two [...] [because] there is no sense to differentiate between the virological load this high or this high, or is it a cirrhotic patient or a non-cirrhotic patient which is [...] not very clearly differentiable in practice. [...] many clinicians said, in theory this is true from the formality of the license, but in practice this can hardly be demarcated whether a patient already has a cirrhosis or not [...] Hepatitis C develops over 30 years and the state of the liver deteriorates gradually and when exactly a cirrhosis is reached [...] that is not easily demarcated. [...] the other point was that IQWiG found a difference in the data according to virological load [...] the FJC said despite seeing this from the data, in practice the virological load always varies [...] [it is] a chance measurement on the basis of which we would be denying a patient treatment [...] and one did [...] not want that (Interviewee No. 5, 2013, p. 3)<sup>31</sup>.

Contrary to IQWiG the FJC decided that it would consider only two patient groups, namely the treatment-naïve group and the treatment-experienced group (G-BA,

---

<sup>31</sup> Translation provided by the author of this thesis.

2012a). The other subgroups that IQWiG highlighted could be grouped within those two groups according to the FJC. While the FJC did not follow IQWiG's complex subdivisions, it also did not feel able to judge with certainty exactly how big the extent of the added benefit of Telaprevir is because the two patient groups for which Telaprevir is licensed includes groups for whom sufficient clinical trial data was not available, thereby rendering the true effect of Telaprevir non-quantifiable (G-BA, 2012a). This suggests that, even though IQWiG's sub-division was not followed, the FJC and IQWiG operate within a decision-making paradigm in which the license for a pharmaceutical product carries significant weight and will determine which subgroups are deemed to be appropriate for the analysis of treatment effects.

The appraisal documentations on early benefits assessments and the interviews conducted as part of this research suggest that, along with concerns about the appropriate comparators for a novel pharmaceutical, the issue around whether the pharmaceutical manufacturers' evidence reflects treatment effects for all possible patient groups for which there is approval status is one that preoccupies much of the work the FJC and IQWiG do. This in itself is not unlike the preoccupations of other HTA institutions because, after all, the market authorisation dictates the legal realm in which a product is allowed to be used and the 'right' comparator product gives a benefit assessment its meaning, i.e. is the use of the new product better, worse or the same than the current treatment pathway. However, the fact that IQWiG and the FJC are comparatively strict when it comes to one of their rules of evidence, i.e. that RCT data has to be available for every possible indication of a new product, has been criticised by stakeholders for its rigidity. The fact that IQWiG and the FJC require convincing evidence for each possible patient subgroup as indicated by the license is reflected in their undertaking to assign an individual, and frequently different, comparator to each of these subgroups that arise from a medicine's approval status. A representative of the pharmaceutical industry criticises the rigorous divisions of the patient populations into subgroups in the following way:

One tries to divide the patient population into [...] subgroups with the result that one cannot serve all of these subgroups on the basis of one's studies [...] post-hoc sub-populations definitions are also not clean methodologically because the study was not powered for it which means one runs the risk that if you, for example, divide your entire population into four sub-populations [...] one might get a significant result for one of them and not for the other three [...] in the end [...] one creates statistical artefacts and there is a risk that one also creates

constellations that are not real anymore and have nothing to do with the [clinical] provision [...] in Germany [...] (Interviewee No. 6, 2013, pp. 4-5).

While subgroup analyses are a part of most benefit assessments they are not without challenges. In addition to methodological challenges there are context-specific challenges that arise from the way HTA institutions operationalise subgroup analyses. Two such context-specific challenges are reflected in the above discussion of IQWiG's and the FJC's approach to subgroup analyses. One is connected to the way a HTA institutions deals with the situation when the evidence presented as subgroup analyses does not yield statistically significant results or when the studies in question do not permit a certain subgroup analysis that is required by the HTA body. Instead of looking to other evidence, IQWiG's and the FJC's 'default' position on such cases seems to be to label the added benefit of a product as 'non-quantifiable' due to lack of scientific evidence.

The categorisation of 'non-quantifiable' added benefit gives rise to the second, and arguably even more pressing, issue, namely the issue of how a certain operationalisation of subgroup divisions will impact on service provision in health care. As the FJC assigns an individual benefit category to each indication for which a pharmaceutical product is licensed, we observe decisions in which a product is classified with different benefit categories for each possible patient subgroup. While the HTA decision-making framework in Germany requires this, it is clear that this might make it difficult for practitioners, payers and patients alike to assess the true value of a new medicine. The benefit category 'non-quantifiable' further complicates these situations as it provides little direction for stakeholders on what benefit to expect from a medicine in comparison to current treatments.

With reference to the question posed in the previous chapter on what constitutes evidence, i.e. what do decision-makers do when the evidence does not yield the answers they seek, it appears that the approach in Germany is to go back to the drawing board in such cases. This means that if the evidence presented is not satisfactory IQWiG will state that this is the case and it will state which data analyses would be needed to give rise to the appropriate evidence. Even though the legal framework in Germany dictates that when evidence from the highest ranking is not available, IQWiG and the FJC should consider evidence from the next level of the ranking (G-BA, 2013), there seems to be an assumption that if the pressing questions cannot be answered by RCT data then

the evidence from the next level of the ranking will also not yield satisfactory answers. This leads to a situation that one stakeholder described as the following:

[...] a premise of evidence-based medicine is ‘do not dismiss available evidence’. Here [in Germany] even evidence from the licensing studies and from other sources that are not in direct relation with the appropriate comparator [...] are just ignored. [...] Purpose of an early benefit assessment so that all those involved gain a certain benefit from it [...] are pragmatic solutions [...] and not to throw available evidence overboard from one day to another due to a paradigm shift [...] (Interviewee No. 6, 2013, pp. 3-4).

The empirical evidence suggests that NICE also considered the question of treatment effects in the various possible subgroups in the case of Telaprevir. A representative of a professional association involved in the appraisal process on Telaprevir stated that “[...] a particular issue was in certain subgroups of people with Hepatitis C, can they have access, what are the risks?” (Interviewee No 3, 2012, p. 1). NICE noted that there were no statistical tests of some of the data in certain subgroups, but when reading its technology appraisal guidance 252 (NICE, 2012e) one gets a sense that this lack of statistical testing was not believed to undermine the overall positive results that the RCTs and the evidence by clinicians and patient experts yielded. The following excerpt from NICE’s guidance illustrates how the Appraisal Committee considered the issue of treatment effects across subgroups:

The Committee discussed the clinical effectiveness of telaprevir plus peginterferon alfa and ribavirin compared with peginterferin alfa and ribavirin alone in previously untreated patients. It noted that the telaprevir-containing regimen statistically significantly increased the sustained virological response rates for ‘standard’ treatment (48 weeks) and response-guided regimens. The Committee observed that telaprevir did not appear to be less effective in patients with cirrhosis than on patients with lower degrees of fibrosis, although it had not been presented with any statistical tests of these data [...]. The Committee discussed the clinical data for telaprevir in patients who had previously been treated. It noted that telaprevir statistically significantly increased sustained virological response rates [...] and that the higher rates [...] were also seen in the patient subgroups (patients whose condition had relapsed, partially responded or not previously responded). [...] It observed that there was no difference in the proportion of patients with cirrhosis who had a sustained virological response to telaprevir compared with the overall trial population (NICE, 2012e, pp. 22-23).

Whereas NICE was satisfied that the higher SVR rates were seen in the patient subgroups of previously treated patients, IQWiG came to a different assessment in that only an indication of an added benefit for non-responders with or without cirrhosis could be derived from the data (IQWiG, 2012c, p. 4). However, for relapsed patients without cirrhosis IQWiG concluded “there were no data on the current approval status in this populations available for the present benefit assessment” (IQWiG, 2012c, p. 6).

The above example shows that subgroup analyses play a role in the appraisal and assessment processes in Germany and in England. However, it indicates a contrasting approach when it comes to operationalising these subgroup analyses when determining the final outcome. While IQWiG and the FJC require data to be translated into statistical test results for every indication of a medicine’s approval status, NICE appears to be willing to assume a positive treatment effect even when statistical tests have not been specifically presented. Presumably this willingness is only present in cases where the overall clinical and cost effectiveness evidence is strong and there are few safety concerns around the pharmaceutical product. Quoting an interviewee who was involved in NICE’s assessment process on Telaprevir:

[...] I think this was a bit of a non-brainer for NICE because the effectiveness of the drug was so much more so the only way that they could fall over was if they killed people [...] or if they priced it so ridiculously that the cost per QALY was outside NICE’s range [...] I think the evidence base was [...] strong (Interviewee No. 1, 2012, pp. 1-2).

The difference in operationalising subgroup analyses is a noteworthy one as it provides an insight into the ‘rules of evidence’ that guide decision-making in Germany and in England. It seems that the availability of statistically significant test results for every patient group that arises out of a pharmaceutical product’s approval status is a precondition for IQWiG to assign the highest benefit category to the product, notwithstanding the possibility that the FJC may overrule IQWiG’s approach in cases such as Telaprevir’s. In contrast, NICE is more open to the idea of assuming effectiveness across different subgroups unless the clinical evidence clearly indicates otherwise or there are serious safety or cost effectiveness concerns. If the evidence is convincing and of good quality overall, NICE appears willing to consider other evidence alongside the evidence from clinical trials in order to make its judgements,

something which played a big role in Telaprevir's case as will be explained in later parts of this thesis.

This aforementioned willingness to consider other forms of evidence may, in part, be due to the fact that according to the NICE's Guide to Social Value Judgements (NICE, 2008a), NICE (2008a) has to specifically justify the in- or exclusion of a patient group in a recommendation in order to avoid discrimination of patients. In analogy to the principle that underlies jurisprudence in democratic states, i.e. 'innocent until proven guilty', the empirical research of this thesis suggests that NICE seems to adhere to a general principle of "effective unless proven ineffective", whereas IQWiG's approach indicates a presumption of "ineffective until proven effective", the possibility that the FJC might overrule some of IQWiG's rigor in assessing treatment effects in patient subgroups notwithstanding. Having said this, it is important to bear in mind that this analogy only pertains to the way the above institutions operationalise subgroup analyses and not clinical or cost effectiveness more generally.

### **7.5.2. Research Findings**

Despite a very different assessment of the strengths and weaknesses of the evidence in the case of Telaprevir and despite the application of different decision-making criteria, the outcome of the benefit assessment was similar in England and Germany. This illustrates that the same outcomes can be arrived at by different means and through different reasoning processes.

The overall appraisal of Telaprevir was positive despite very different takes on the appropriateness of the data and the criteria applied to it. What remains constant and similar to the cases presented earlier in this chapter is that the interpretation of the evidence, i.e. the criteria, standards and algorithms applied to it are different in Germany and England. However, that is not to say that these diverging criteria necessarily result in diverging outcomes. The case of Telaprevir illustrates that similar outcomes can be arrived at by different means, i.e. despite different inputs and interpretations. As well as offering interesting insights on what determines the outcome of pharmaceutical benefit assessment, the discussion on Telaprevir also underscores the merits of process-tracing exercises when analysing complex processes. Without tracing the questions that were addressed by NICE, IQWiG and the FJC in this case, one might have been tempted to infer a similarity of argumentation from the similarity in the

dependent variable, which would not have represented an accurate account of what determined the outcome in this case.

In summary, the discussion on Telaprevir leads to the following research findings:

1. The case of Telaprevir exemplifies that contrasting HTA policy paradigms do not necessarily lead to different outcomes in the dependent variable. This finding is significant because it arises from a case scenario where logic would have led one to believe that the outcome would be different (i.e. different input + different criteria = different outcome, instead we saw different input + different criteria = similar outcome). This shows that one cannot assume that a similar outcome in the dependent variable is the result of undisputed and strong evidence. Even cases in which the outcome in the dependent variable is similar, different rules of evidence are applied that help explain this outcome.
2. Similarly to the previous sections on Fingolimod and Retigabine, the discussion of Telaprevir highlighted the crucial role of the decision-making criteria, the rules of evidence, in determining the outcome of the benefit assessment.
3. NICE's paradigms appears to be more secure than the German HTA paradigm. The contrast between NICE's decision-making ease and the FJC's and IQWiG's decision-making struggles in this case are striking. Whilst NICE dealt with the issues of the case with comparative ease, IQWiG and the FJC struggled with the articulation and application of their paradigm in the case of Telaprevir. Both the limited interpretation of patient relevance as well as the strict application of the market authorisation by the FJC would have led to outcomes that would have been considered unacceptable and unreasonable by stakeholders. However, this finding should not be taken at face value as it is based on one case only and needs to be supplemented by evidence from other cases that were examined in this thesis.
4. The case of Telaprevir provides an important theoretical insight on how contrasting policy paradigms do not, as a matter of course, lead to divergent outcomes in the dependent variable.

In the final section of this chapter I discuss the fourth theme that emerges from the empirical evidence, namely the question of the operationalisation of thresholds and algorithms under HTA paradigms.

## **7.6. Theme Four: Operationalisation of Thresholds and Algorithms under HTA Paradigms**

In addition to the themes discussed in the previous sections, another subject that played a role in HTA decision-making processes was that of the operationalisation of the HTA decision-making paradigm by way of thresholds and algorithms. In addition to the definition of the decision-making criteria such as cost effectiveness or patient relevance, these criteria might be operationalised by placing numerical and statistical values to them above or below which they are satisfied. For example, NICE's guide to technology appraisals outlines a threshold range of an ICER of £20,000-£30,000/QALY as one that will usually be deemed a cost effective use of NHS resources (NICE, 2008).

The question of thresholds and algorithms is connected to this thesis' research question because, when they exist, they potentially provide a straightforward answer to what determines the outcome of pharmaceutical benefit assessments. That is, in the presence of decision-making algorithms in the form of numerical and statistical thresholds that have to be met, stakeholders are aware of what a new product needs to achieve with regards to clinical or cost effectiveness. Expressing the value of a medicine in a numerical fashion provides an idea of what will be considered beneficial in an HTA process. However, despite the apparent advantages of HTA algorithms in the form of cost effectiveness thresholds, statistical expressions of value also carry the danger of under- or over-valuing certain aspects of a new treatment that cannot easily be expressed in numbers. In Germany and England we observe two cases in which decision-making algorithms are present, albeit in different formats. As will be shown in the following paragraphs these algorithms influence the outcome of pharmaceutical benefit assessments in both countries.

Previous sections outlined that there are exceptional situations in which the recommendation of a medicine above NICE's cost effectiveness range between £20,000-£30,000/QALY might be acceptable, but as a general rule, the threshold range applies. In terms of NICE's cost effectiveness paradigm this means that a clear decision-making algorithm exists. A comparable algorithm for the benefit categories is, thus far, lacking in Germany.

In an appendix to its first early benefit assessment in January 2011, IQWiG laid out its methods for the operationalisation of the extent of added benefit (IQWiG, 2011). According to IQWiG this was necessary because even though the rules of procedure for

early benefit assessment provided some guidance by way of definition of the most important criteria for the different benefit categories, the effect variables still had to be set up in an hierarchical manner according to their importance and decisions had to be made about the desirable effect sizes the variables had to meet in order to be considered proof of a certain category (IQWiG, 2011). The appendix to IQWiG's first early benefit assessment, also referred to as the 'Ticagrelor appendix' (named after the pharmaceutical product that was assessed), outlines the statistical thresholds in relation to confidence intervals and relative risk ratios that a product needs to meet or overcome in order to be assigned to a certain benefit category (IQWiG, 2011, pp. 86-92). Since January 2011, this appendix has served as the basis for IQWiG's operationalisation of the extent of added benefit.

The Ticagrelor appendix and the operationalisation of added benefit has received much criticism by stakeholders in the processes of early benefit assessments. The criticisms centre on the lack of scientific validation of the method that IQWiG uses and the lack of transparency in relation to whether the FJC follows this operationalisation or not. In all of the cases that were examined as part of this thesis, the FJC underlines that it did not adopt IQWiG's method for the operationalisation for added benefit. The phrasing for this iteration is near-to identical in all cases and reads, by way of example, in the following way: "The method proposed by IQWiG in appendix A of the benefit assessment dossier on Ticagrelor [...] was not relied on in the benefit assessment of Apixaban"<sup>32</sup> (G-BA, 2012b, p. 3). However, despite this, the instances in which the FJC does not follow IQWiG in its assessment of the added benefit are comparatively rare. In the cases that were analysed as part of this thesis these instances include Boceprevir and Telaprevir in which the FJC came to a different conclusion on the question whether the main clinical endpoint of sustained virological response (SVR) rate was patient relevant or not. Other than that, the FJC followed IQWiG's benefit categorisations, thus putting a question mark on its statements that IQWiG's methods are not followed by the FJC.

The most common criticism by stakeholders during the hearing process at the FJC centers around the lack of clarity and transparency on how the added is benefit is operationalised by the FJC (e.g. G-BA, 2012c; 2012g). Furthermore stakeholders have criticised IQWiG's methods for the operationalisation of added benefit as scientifically

---

<sup>32</sup> Translation provided by the author of this thesis.

unfounded and lacking a scientifically substantiated definition (G-BA, 2012c, p. 121). A representative of the pharmaceutical industry phrased the issue in the following way: “My impression is [...] that no algorithm exists. It’s decisions made on individual cases which does not necessarily make the situation easier because we cannot draw lessons from it” (Interview No. 6, 2013, p. 5). This perception of the FJC’s operationalisation of added benefit is underlined by a representative of the HTA decision-making body in Germany:

[...] I can’t say that [...] the FJC has developed a matrix for making its decisions. That’s just not the case until now [...] I also don’t know whether our system is earmarked for this, for structuring something in such a mathematical way that it becomes predictable [...] it won’t become that way (Interviewee No. 16, 2013, p. 2).

Even though the FJC reiterates in most benefit appraisal documents that it has not adopted IQWiG’s method of the operationalisation of added benefit, the fact that the FJC follows IQWiG’s categorisations more often than not is a good indicator that IQWiG’s Ticagrelor appendix serves as the an algorithm-like matrix within the German HTA paradigm. An interviewee supports this view in the following way:

[...] the FJC follows [IQWiG] most of the time, the FJC says that it does not use our methods but it follows our assessments most of the time. So either they are using our methods after all, then they [the methods] cannot be that bad or the FJC is conducting an independent study to show validity, then the methods can also not be that bad (Interviewee No. 23, 2013, p. 4).

To summarise, the HTA decision-making algorithms in Germany and England differ in both substance and explicitness. They differ in substance in that the English algorithm is based on cost effectiveness ratios whereas the German algorithm is based on statistical thresholds that patient relevant endpoints have to meet in order for a medicine to be assigned to a certain benefit category. They differ in explicitness in that the English paradigm is explicit with regards to necessary thresholds and algorithms by including them in the relevant methods guidelines. Germany’s HTA is less explicit about its paradigm which gives rise to some confusion and controversy during the early benefit assessment processes. While IQWiG has developed an algorithm in the form of the Ticagrelor appendix, which will be incorporated in a new version of IQWiG’s methods guidelines in 2014 (Interviewee No. 23, 2013), the FJC is more reluctant to

subscribe to a decision-making algorithm. The empirical evidence does not give rise to conclusive reasons for this and possible reasons that I could offer here would run the risk of being speculative in character. Suffice it to say that, regardless of the FJC's apparent reluctance to adopt an algorithm, the empirical evidence suggests that it adopts IQWiG's methods implicitly when it follows IQWiG's benefit categorisations of new medicines.

### **7.6.1. Research Findings**

When putting the question of thresholds and algorithms into the context of the previously discussed themes in the empirical data, we can infer that the presence or absence and the operationalisation of algorithms play an important role in determining the final outcome of pharmaceutical benefit assessments. This is because once the HTA decision-makers have decided which evidence to use, which comparator to use and which patient populations to analyse, the results pertaining to these parameters are measured against the qualitative and quantitative standards that the respective algorithms prescribe. Whether or not these standards are met and whether, in case they are not met, exceptions can be made influences the outcome of a benefit assessment. For example, in the previously discussed examples of Cabazitaxel and Eribulin (see tables 6.3. and 6.4.) the results on cost effectiveness and clinical effectiveness met neither the standards of NICE's cost effectiveness paradigm nor the standards of the FJC's patient relevant paradigm. The operationalisation of the HTA paradigms in terms of thresholds and algorithms can thus be described as providing the tool for an assessment and appraisal, whereas the previously mentioned themes make up the substance that the tool gets applied to. The above analyses highlight that both the substance and the tools matter when it comes to determining the outcome of pharmaceutical benefit assessments.

In conclusion, the brief discussion on the operationalisation of HTA paradigms contributes to the following research findings:

1. The discussion reaffirms that the rules of evidence matter. The discussion on the operationalisation of thresholds and algorithms demonstrates that they especially matter when they pertain to what Hall (1989) refers to as the instrument settings of a policy. In the case of HTA policy these instrument settings are the algorithms that are applied to the scientific data in order to interpret it. Different

instruments settings lead to different interpretations of the evidence, thus supporting the finding that the ‘rules of evidence’ are the most important variable in determining the outcome of pharmaceutical benefit assessments in systems that employ formalised HTA procedures.

2. The discussion also supports one of the preliminary conclusions drawn in chapter 5, which put forward the possibility that the operationalisation and conceptualisation of thresholds play a significant role in determining the outcome of pharmaceutical benefit assessments as notable differences emerged from the analysis of the methodological guidelines. The empirical evidence in this chapter further supports the conclusion that the operationalisation of thresholds and algorithms under different paradigms matters to the final outcome.

## **7.7. Conclusion**

This chapter discussed the decision-making considerations in the cases of Fingolimod, Retigabine and Telaprevir. It showed that in addition to the question of what constitutes evidence in the first place (theme one) the way in which this evidence is addressed, or interpreted, plays an important role in determining the outcome of pharmaceutical benefit assessments. The appropriateness of the choice of comparator (theme two) and the appropriateness of the sub-division of patient populations (theme three) are important components in the process of making the evidence relevant to the given HTA context. The prominence with which they were discussed by NICE, the FJC and IQWiG in the aforementioned cases suggests a similar importance of these variables in both the English and the German context. However, the substantive differences with which they were approached suggest that different rules of evidence exist. However, a difference in rules of evidence and a difference in assessment of what is considered appropriate evidence should not be equated with an inevitable difference in the assessment outcome. As the case of Telaprevir shows the combination of different factors and reasoning can ultimately lead to the same outcome, thus suggesting, at least for HTAs, that there is more than one causal path leading to the particular score in the dependent variable.

Even though theme four was discussed comparatively briefly in this chapter, its significance in determining the outcome of pharmaceutical benefit assessments should

not be underestimated. No matter what other questions certain pieces of evidence give rise to and how these questions are resolved, in every case the ultimate hurdle that needs to be passed in HTAs is that of thresholds. Whether thresholds are set formally with the help of ICERs or more generally with the help of benefit categories, they are an expression of the limit of the price a society is willing to pay for a pharmaceutical product and this limit depends on how clinical benefits are conceptualised and measured (see discussion in chapter 5).

Since thresholds set the limits, or boundaries, of what is considered acceptable they are one of the most important elements of any HTA paradigm. They function in much of the same way that paradigms function. To recall, paradigms function as intellectual constructs that determine what is or is not possible (chapter 2); thresholds have a similar boundary-defining role in determining the limits for benefit assessments. It is therefore not surprising that one of the conclusions I draw in this thesis is that the dominant HTA paradigms in England and Germany can be described according to the component that relates to their thresholds, i.e. cost effectiveness in England and patient relevance in Germany. Thresholds and algorithms are a crucial factor to understanding what determines pharmaceutical benefit assessments.

In the next chapter I outline the sixth and final theme that was raised in the empirical evidence. This theme relates to auxiliary variables that might play a contributory role in determining the outcome of pharmaceutical benefit assessments.

## Chapter 8

### **Policy Paradigms in Operation III: Politics and the Articulation of Paradigms**

#### **8.0. Introduction**

In this final chapter of the empirical analysis I discuss an additional theme (table 6.1, theme six) that may contribute to the outcome of pharmaceutical benefit assessments in England and Germany. In contrast to the themes discussed in the previous two chapters, the issues presented here did not arise in every embedded case study that was examined. This situation warranted a separate chapter for the discussion of these topics.

The theme discussed here is summarised in table 6.1. under the theme heading ‘question of political power and influence’<sup>33</sup>. Similarly to the previous themes, this theme have rise to different concerns and foci in England and in Germany. These concerns represent the sub-themes that emerged from theme six and are referred to as auxiliary variables in this chapter. The variables are labeled ‘auxiliary’ because they do not appear to contribute to the final outcome of the pharmaceutical benefit assessment as a matter of course, that is in they do not matter in each and every case. Whereas the previous themes around evidence questions, the appropriate comparator and subgroups of patient populations were crucial elements in every benefit assessment, the variables of public pressure, stakeholder bargaining power and controversies around key paradigmatic issues only featured in a couple of cases. In the interest of comprehensiveness and of highlighting areas for future research the aforementioned topics are discussed in this chapter, but the discussion highlights that the role they play in contributing to the outcome of pharmaceutical benefit assessments requires further exploration that is beyond the scope of this thesis.

The chapter proceeds by providing an overview of the different auxiliary variables that arise from the empirical evidence in England and Germany. The auxiliary variable that arises from the evidence in England is that of ‘public pressure’. The

---

<sup>33</sup> The term ‘political’ is used in its broadest sense. That is to say that political power and influence is not restricted to politicians and policy-makers, but includes all political actors (e.g. lobbying groups, charity organisations, pharmaceutical industry, doctors etc.) who have a stake in the outcome of HTAs.

evidence suggests that this variable potentially contributes to the outcome of pharmaceutical benefit assessments in England when stakeholder groups make a decision to get involved because they believe that the clinical effectiveness data on a product warrants stakeholder opposition to a decision made by NICE. For Germany, the auxiliary variables include the bargaining power of certain stakeholders vis a vis others as well as the conceptualisation of the paradigmatic idea of ‘patient relevance’. The section on Germany highlights that, while it is likely that the second of these variables contributes to the outcome of the dependent variable, the role of the first one is slightly more caveated, due to reasons of potential bias and unavailability of data that will be outlined.

### **8.1. Theme Six: Auxiliary Variables Emerging from the Empirical Evidence**

The previous chapters provided a discussion and analysis of the five main themes that the empirical evidence gave rise to. These themes included questions about what constitutes evidence in the first place, what the appropriate comparator and patient sub-populations are, how thresholds are operationalised and how the decision-making paradigms address situations in which RCT evidence poses challenges.

The analysis showed that while the five themes affect the outcome of pharmaceutical benefit assessments in Germany and in England, the way in which the themes are addressed varies according to the policy paradigms that are applied. Whereas the significance of the ‘rules of evidence’ was unequivocal in Germany and England, the empirical evidence suggests that additional factors, or auxiliary variables, are at play in both countries. In England the additional factor that played a role is the question of how public pressure and criticisms might contribute to the outcome of certain cases. In Germany, the additional factors are the relative bargaining power of the federal association of statutory sickness funds vis a vis other stakeholders and the conceptualisation and operationalisation of the idea of patient relevance. Due to the political character of these variables they can be subsumed under questions of political power and influence, thus suggesting a connection to the political dimensions of HTA that were outlined in chapter 1.

In the interest of comprehensiveness, the aforementioned topics are discussed in the following sections. However, because they represent outliers in the data that was collected – that is to say they were mentioned in some, but not all of the embedded case

studies – the analysis of the topics will be comparatively brief. I hope that including them in this thesis opens future avenues for exploring the role that public pressure and bargaining power play in determining the outcome of pharmaceutical benefit assessments.

## **8.2. Theme Six: Public Pressure as an Auxiliary Variable in England**

The nature of HTA processes, especially when they have an impact on the availability of new pharmaceutical products within a given health care system, is such that HTA bodies frequently face criticism in the public arena, which may come in the form of media campaigns, protests headed by patient representatives or clinicians, and/or involvement of policy-makers such as politicians. This trend continues despite the hope that HTA processes might contribute to de-politicising the challenging decisions of pharmaceutical coverage. In NICE's case, the empirical evidence suggests that the issue of public pressure may have played a role in two, namely Abiraterone and Fingolimod, out of the ten cases of pharmaceutical benefit assessments that were studied. However, the relative effect of public pressure in the cases of Abiraterone and Fingolimod is difficult to measure as it is not a variable whose significance is easily admitted by stakeholders. Nevertheless, an overview of how the public became involved in the two cases suggests that it may have been an auxiliary variable that contributed to the outcome in these cases.

In the case of Fingolimod (see table 7.1.) the public pressure on NICE arose in the form of two letters by Members of Parliament (MPs). One letter made reference the problem of the 'postcode lottery' and the comparatively poor performance of the UK in terms of access to new treatments for multiple sclerosis. The MP in question wrote:

I am [...] very concerned that a final NICE decision not to approve this medication would leave some people with MS unable to access an effective treatment option, thus exacerbating both the 'postcode lottery' of MS treatment that already exists in this country and the relatively poor approach taken by the UK to MS care when compared with other European countries (McKinnel, 2011).

The fact that the postcode lottery and its associated issues of inequality and discrimination are referred to in the above quote is interesting as it reflects an

appreciation of the values contained in the wider English health care paradigm. As outlined in chapter 5 equality, non-discrimination and an end to the postcode lottery are amongst the core values that NICE's paradigm is built on. The above quote suggests that stakeholders in England are aware of this and that this awareness might extend to a belief that NICE might re-think its position if one appeals to its core values of equality and non-discrimination.

The MPs expressed their concerns as part of the consultation process at NICE. As the consultation process is open to the public, this in itself would not be noteworthy. However, when compared with the other nine cases of pharmaceutical benefit assessments, Fingolimod is the only case in which MPs got involved in the consultation process. Unfortunately, while this is a noteworthy and interesting observation, it does not permit any meaningful inferences about the impact this had on the outcome of the decision in this case, especially because it was not accompanied by a significant public campaign in the media or otherwise. Ultimately, and according to the guidance produced by NICE (2012d), the fact that the pharmaceutical manufacturer of Fingolimod handed in a patient access scheme for the product seems to have contributed more significantly to NICE's ability to recommend Fingolimod than the fact that two MPs got involved. However, by the same token, the influence of the MPs' involvement as an auxiliary variable cannot be entirely ruled out on the basis of the empirical evidence.

The empirical data in the case of Abiraterone (table 8.1.) seems to provide stronger evidence than the data pertaining to Fingolimod to suggest that public pressure may have influenced the final decision by NICE. Abiraterone (NICE, 2012f) is a pharmaceutical product that is licensed for castration-resistant metastatic prostate cancer previously treated with specific courses of chemotherapy. Abiraterone's mode of action is novel in that it is not a chemotherapy, but an anti-hormonal treatment, which is associated with less side effects than chemotherapies (NICE, 2012f).

**TABLE 8.1. – Case Study: Abiraterone (NICE, 2012f; G-BA, 2012h; IQWiG, 2011c)**

Indicated for: Castration-resistant metastatic prostate cancer previously treated with a docetaxel-containing regimen

	NICE (2012)	FJC (2012)
<i>Dependent variable: Outcome of benefit assessment</i>	Recommended ( <b>Positive outcome</b> )	Differing benefit assessment for 2 distinct patient populations ( <b>Positive outcome overall</b> ): - For ‘best-supportive’ care population: Indication for a significant added benefit - For Docetaxel re-therapy population: Added benefit not substantiated
<i>Reasoning/discussions/topic raised/public context</i>	<ul style="list-style-type: none"> <li>- Initially not recommended, only recommended after presentation of new evidence</li> <li>- Patient Access Scheme (PAS)</li> <li>- End-of-life criteria met (initially not met because eligible patient population was considered too large)</li> <li>- Considered a ‘step change’</li> <li>- Benefit of oral treatment not captured in QALY</li> <li>- Significant public campaign following the initial decision of NICE not to recommend it</li> <li>- Evidence issues: Patient population by manufacturer did not match the marketing license</li> </ul>	<ul style="list-style-type: none"> <li>- Only accepted one out of four randomised controlled trials (RCTs) because only one matched patient population and comparator as specified by FJC</li> <li>- FJC determined additional patient relevant endpoints</li> <li>- Substantial critique by stakeholders of second subgroup division as ‘docetaxel-retherapy’ as a treatment option is not deemed to be underlined by sufficient evidence</li> </ul>
<i>Differences/similarities</i>	<p>→ <b>Positive recommendation, but only on second attempt</b></p> <ul style="list-style-type: none"> <li>- Accepted four RCTs</li> <li>- Specification of clinical outcomes accepted as presented by manufacturer</li> <li>- Subgroup divisions accepted</li> <li>- Saw problems with lack of congruence between RCT population and licensing population</li> <li>- Questions around clinical practice</li> <li>- Accepted ‘oral treatment’ characteristic as a point for consideration</li> </ul>	<p>→ <b>Positive recommendation, highest benefit category but only for one patient subgroup</b></p> <ul style="list-style-type: none"> <li>- Accepted one RCT</li> <li>- Additional patient relevant endpoints required in addition to outcomes laid out by manufacturer</li> <li>- Additional/different subgroup divisions made</li> <li>- Saw problems with lack of congruence between RCT population and licensing population</li> <li>- Questions around clinical practice</li> <li>- Did not accept ‘oral treatment’ characteristic as a point for consideration</li> </ul>

Table 8.1. shows that NICE and the FJC made an overall positive decision for Abiraterone. However, in NICE's case this decision was only arrived at after reversing the draft recommendation, which did not recommend Abiraterone on the grounds of being cost ineffective. The FJC categorised the additional benefit of Abiraterone differently for the two relevant patient populations. It assigned the category of 'significant added benefit' to the patient population who are no longer eligible for another course of chemotherapy and the category of 'added benefit not substantiated' for those who are still eligible for another course of chemotherapy (G-BA, 2012h). Besides Ipilimumab, Abiraterone represents the only case included in this thesis in which the FJC assigned the category of 'significant added benefit', which is why, overall, the FJC's decision can be described as positive in this case. Against the backdrop of the discussion on what constitutes evidence that was presented in chapter 6, it is worth highlighting that NICE (2012f) accepted four RCTs as the evidence base in the case of Abiraterone, whereas the FJC only accepted one trial that, according to the FJC, matched the licensed population and appropriate comparator, thereby highlighting once again the importance of theme one and two (G-BA, 2012h).

In its Appraisal Consultation Document (ACD), a document that is published if the proposed recommendation is negative in order to allow for a consultation process, NICE did not recommend Abiraterone (NICE, 2012g). This was because the ICERs presented by the manufacturer – and that were already high at £63,200 per QALY – would likely increase if the utility values of different health states were modeled differently in the health economic model (NICE, 2012g). The initial recommendation by NICE was met with large protests in the form of an extensive media campaign by patient groups. The campaign criticised NICE's draft recommendation on a number of issues that, broadly speaking, sought to appeal to the values base of the work that NICE does. Most importantly, it criticised that NICE was not convinced that Abiraterone met NICE's end-of-life criteria because the eligible patient population was not considered to be small.

As part of the public campaign that followed NICE's publication of the ACD, statements in the media included, but were not limited, the following assertions about the draft recommendation:

Cancer Research UK said the draft decision by the National Institute for Health and Clinical Excellence [...] made "no sense" and that Nice had used the wrong

criteria to judge the drug. [...] Owen Sharp, chief executive of the Prostate Cancer Charity, said: “[...] This draft is a bitter blow to thousands of men and their families and must be overturned” (Boseley, 2012).

Abiraterone is the latest prostate cancer drug to face an NHS ban despite being proven to extend life for men with advanced disease [...]. It is a fresh blow for doctors and patients who hoped a new era of drugs could lessen the deadly toll of prostate cancer, which has been described in the past as a low-profile ‘Cinderella’ disease (Hope, 2012)

These examples highlight the protests that NICE faced around its draft recommendation. Eventually, NICE reversed its original position and recommended Abiraterone as a treatment option. This was possible because it accepted that Abiraterone fulfilled the end-of-life criteria, which it had not originally, and because it was presented with new health economic analyses (NICE, 2012f). Explaining the reversal of the original recommendation, Sir Andrew Dillon, chief executive of NICE, was quoted in a major media outlet as saying:

During the consultation on the draft guidance [...], the manufacturer of the drug, submitted further information for the committee to consider. This included a revised patient access scheme which involves providing the drug to the NHS at a discounted price, further information on which patients would benefit the most and clarification on how many patients could receive the drug. These factors enabled the committee to revise its preliminary recommendation and now recommend the drug for use on the NHS (Edgar, 2012).

While the above statement by Sir Andrew does not include a reference to the public campaign that preceded NICE’s preliminary recommendations, a representative of a patient charity that was involved in the public campaign on Abiraterone suggested that the campaign generally, and the charity’s involvement in it specifically, may have contributed to NICE’s reversal of its position:

We did do some work publicly, and also behind the scenes, urging the manufacturer to reduce their price in order for Abiraterone to be made available [...] I think that’s important and leading on from there [...] is the [...] public campaigning [...] so even though I think the committee are quite rigid [...] in terms of sticking to their guns [...] if they don’t see the evidence and they don’t see something as being cost effective then they will mostly still continue to say ‘no’ but I think what we did do was to communicate to our supporters and also to the media the case around Abiraterone and why it should be made available [...] we did make a difference there [...] there was the public and there was the behind the scenes [...] in terms of an argument and in terms of influencing [...]

counteracting what NICE was saying in their draft decision and reasons why we didn't think it was correct [...] it made them think again [...] (Interviewee No. 19, 2013, p. 2)

The above quote reaffirms the strength of the media campaign around Abiraterone, and the possible contribution it may have had to the final outcome, but it also raises an additional point, which suggests that Abiraterone might be an exceptional case. That is the point about the evidence and cost effectiveness. In Abiraterone's case the clinical evidence was very strong and the stakeholders agreed that is was. By contrast, in Cabazitaxel's case, which is licensed for almost the identical patient population as Abiraterone is and which was appraised in the same year as Abiraterone, the clinical evidence was less convincing and the Appraisal Committee did not reverse its decision to not recommend it. Even the charity that was involved in the consultation process in both cases admitted that there were flaws in the evidence for Cabazitaxel:

[...] we were of the view that in terms of patient data and in terms of safety and quality of life issues we didn't feel as if the manufacturer had provided enough evidence or the trial hadn't shone a light on those concerns [...] we were disappointed that NICE didn't recommend Cabazitaxel but we had our own concerns [...] we weren't too upset in terms of Cabazitaxel. Abiraterone was something completely different [...] it was clinically effective, also had benefits in terms of quality of life and it was life-extending [...] also Cabazitaxel is administered intravenously whereas Abirateron is in pill-form, so overall it was [...] very much an effective drug (Interviewee No. 19, 2013, p. 1).

The comparison with Cabazitaxel, a product licensed for the same patient population as that of Abiraterone and at the same time, suggests that the success of a public campaign depends on the strength of the evidence of the clinical benefits of a new product. The empirical evidence presented above makes it difficult to envision a situation in which public pressure might lead to a reversal of NICE's decision despite weaknesses in the evidence. The possibility cannot be ruled out, but the above data suggests that convincing evidence, and the appropriate economic modeling thereof, remains the key that opens the possibility for NICE reversing a decision.

It is also important to note that Abiraterone may in itself be an unique case with exceptional features in that it is a) not a chemotherapy despite the alternative treatments for the licensed population consisting mostly of chemotherapies and b) a product that can be administered orally and was thus seen as providing additional benefits by

patients and clinicians alike. In addition to the strong evidence on its clinical effectiveness, these special features of Abiraterone seemed to have contributed to its positive appraisal in the end.

The unique characteristics of Abiraterone and the strong clinical evidence arising from the clinical trials also appear to have contributed to stakeholders such as the main charity for prostate cancer to embark on a big public campaign. This claim arises from the previously presented observation that the charity decided to become active in the case of Abiraterone, but not in the case of Cabazitaxel. This suggests that, at least in the case of this charity, the commencement of a public campaign is weighed against its likely success. Since the evidence on Cabazitaxel was weak and the evidence on Abiraterone was strong it makes sense that the latter product became the centre of an extensive media campaign.

Overall, the consultation documents, stakeholder interviews and media excerpts on Abiraterone suggest that NICE's reversal of its decision was a result of reframing the clinical effectiveness evidence and adjusting the cost effectiveness calculations. Upon being asked about what kind of issues determined whether NICE would change an originally negative recommendation, a former executive director at NICE confirmed the previous interpretation:

[...] the biggest issue was interpretation of effectiveness and cost effectiveness [...] there would be a debate about what efficacy, effectiveness data should be used [...]. The other big problem is that there may not be quality of life data in the trials, so there would be an application of quality of life from some other studies and applied them to the trial [...]. Another big issue was the time horizon [...] it was normally very very – subjective is the wrong word – but depending on your viewpoint, what was the best evidence? (Interviewee No. 11, 2013, p. 4).

Together with the previous statements drawn from stakeholder interviews, the above quote actually reaffirms the significance of the 'rules of evidence', i.e. what constitutes evidence and how it should be interpreted, even when NICE is faced with public pressure or when its decisions are being challenged. This is what the above quoted interviewee referred to as “[...] depending on your viewpoint, what was the best evidence” (Interviewee No. 11, 2013, p. 4). In the end, a strong public campaign might contribute to a reassessment of the original interpretation of the evidence, but the evidence itself is at the core of how NICE justifies its ultimate decision. The presented

empirical analysis suggests that the public campaign on Abiraterone highlighted the need for approaching the evidence from a different angle, but that it was an auxiliary rather than a stand-alone factor that contributed to the final outcome of NICE's appraisal of Abiraterone.

### **8.2.1. Research Findings**

The cases of Fingolimod and Abiraterone suggest that public pressure can act as auxiliary variable that contributes to the outcome of pharmaceutical benefit assessments in England. However, based on the sample of cases that was analysed in this thesis, this appears to be the exception rather than the rule. In order to better understand the role of public pressure, a separate research project that examines the cases in which there was a significant public campaign and whether this led to a reversal of NICE's recommendation is called for. While this is beyond the scope of this thesis, the empirical evidence does suggest that public pressure can play a contributing factor in determining the outcome of pharmaceutical benefit assessments in England.

The above discussion highlights that the empirical evidence gives rise to the auxiliary variable of 'public pressure' in the context of the English HTA paradigm. It should be labeled as auxiliary for two reasons. Firstly, the cases that give rise to a significant public campaign seem to be the exception rather than the rule. The case of Abiraterone suggests that whether a public campaign is initiated depends on the relative strength of the available clinical and cost effectiveness evidence, but more case studies would be required to confirm this interpretation. The second reason for why public pressure should be labeled auxiliary is based on the observation that even in the instances in which a public campaign is initiated and NICE's decision is challenged, the reasoning process used by NICE and the involved stakeholders comes back to question of the 'right' evidence and the appropriate interpretation thereof. The public campaign might contribute to raising the awareness around a product, but ultimately NICE and the involved stakeholders frame their arguments around evidence questions. This in turn reaffirms the significance of the 'rules of evidence' of what determines the outcome of pharmaceutical benefit assessments.

In conclusion, the discussion on the role of public pressure as an auxiliary variable in the English context gives rise to the following research findings:

1. The auxiliary variable in the English HTA paradigm is public pressure. It is auxiliary in the sense that it does not arise in every case and when it does, the ultimate decision-making reasoning by NICE still centres on evidence-based questions.
2. The way in which the public campaign and protests were framed in the case of Fingolimod and Abiraterone confirms the existence of a shared value-based HTA paradigm. In both cases public statements included references to values such as equality, non-discrimination, and a general concern for the fate of patients, all of which are at the heart of the wider paradigm under which the NHS operates. This confirms a) that policy paradigms exist in HTA systems and that they reflect wider ideas and values and b) a consonance between the HTA paradigm as it is laid out in theory and how it is operationalised in practice in England.
3. The fact that arguments and re-interpretations are framed with reference to evidence questions by stakeholders such as patient groups and decision-makers, even when re-assessing a decision amidst public pressure, reaffirms the central role that the 'rules of evidence' play in determining the outcome of pharmaceutical benefit assessments.

For Germany the empirical evidence does not suggest that public pressure plays a role as an auxiliary variable. Instead the questions of bargaining strength amongst stakeholders and the operationalisation of patient relevance arise within the theme pertaining to political power and influence. These questions are discussed in the next section.

### **8.3. Theme Six: Bargaining Powers and the Idea of Patient Relevance as Auxiliary Variables in Germany**

Individual pharmaceutical benefit assessments receive comparatively little attention in the public arena in Germany. There was no significant public campaign in any of the cases that were analysed as part of this thesis. Public pressure as an auxiliary variable in Germany does not seem to play the role it does in England. Instead, the empirical evidence indicates that the relative bargaining power amongst stakeholders, as well as the operationalisation of patient relevance, might contribute to the outcome of pharmaceutical benefit assessments. In relation to the latter, stakeholders perceive a dissonance between the theory and the practice of the concept of patient relevance. In terms of the former, stakeholders are suggesting that the strong position of the federal association of statutory sickness funds in the early benefit assessment process may impact on the ultimate outcome of the assessments in the sense that the association can shape the decision problems of the assessments in a way that benefits their interests. A brief overview of how these two issues are constructed in the empirical evidence is given in the following sections.

As outlined in chapter 4, the institutional structure of the early benefit assessment process in Germany is such that the federal association of statutory sickness funds, that is the umbrella organisation for all sickness funds and the representation of payers' interests, is involved at every stage of the assessment process. As a member of the FJC the federal association of statutory sickness funds is involved in the decision on what the appropriate comparator should be. Upon receiving a recommendation by IQWiG, it is then involved in making a decision on the added benefit of a given pharmaceutical product. Finally, the association's representatives negotiate the reimbursement price based on the early benefit assessment and other considerations in the price negotiations with the pharmaceutical manufacturers. In the empirical data gathered as part of this thesis, the role of the association of statutory sickness funds has been characterised as "dominant". The concern over the strong position of the federal association of statutory sickness funds was especially raised by pharmaceutical representatives:

[...] the question of the appropriate comparator is crucial in the whole process and its connected to another problem, namely that the same people always decide on the crucial procedural steps [...] the FJC dominates through the

federal association of statutory sickness funds [...]. [It] decides on the appropriate comparator, then on the added benefit and in the end it negotiates with the manufacturer (Interviewee No. 14, 2013, p. 3)<sup>34</sup>.

According to the pharmaceutical representatives who were interviewed in Germany the strong position of the federal association of statutory sickness funds impacts on the outcome of pharmaceutical benefit assessments because the choice of the appropriate comparator essentially determines how the assessment will proceed and what the likely outcome will be. As discussed in the case of Retigabine, IQWiG (2012d) and the FJC (2012f) were not of the view that the appropriate comparator was chosen for the early benefit assessment. This decision rendered the manufacturer's submission and the analysis of the clinical effectiveness evidence useless to such an extent that the presented evidence was not even considered by IQWiG and the FJC because it was based on the 'wrong' comparator. The choice for comparator, to which the federal association of statutory sickness funds can contribute within the FJC, had a real impact in this case.

Connected to the perception that the strong position of the federal association of statutory sickness funds plays a big role in the German HTA system is another concern that was raised frequently during the interview process with pharmaceutical manufacturers, namely the concern that the association has a stronger bargaining position in the price negotiations by virtue of the fact that it is involved in every step of the assessment process and that it is an experienced negotiator due to the number of negotiations it is involved in. One representative of the pharmaceutical industry alerted to this issue as follows:

[...] we have a classic monopolisation due to the fact that it [the federal association of statutory sickness funds] always negotiates, it has extreme learning curve effects [...] it has a market dominance [...], it can develop tactics [...] it is an ideal world for the federal association of statutory sickness funds. For the manufacturers the situation is: some have more experience because they have done 2 or 3 negotiations [...] others have one product, small biotechnical manufacturers [...] they have one product every 3 years and they are completely disadvantaged [...]. [...] the federal association of statutory sickness funds will have 120 negotiations in 3 years (Interviewee No. 6, 2013, pp. 9-10)<sup>35</sup>.

---

<sup>34</sup> Translation provided by the author of this thesis.

<sup>35</sup> Translation provided by the author of this thesis.

While the above concern moves beyond the question of what determines the outcome of pharmaceutical benefit assessments as it relates to the stage that comes after the benefit assessment, it is still noteworthy because it highlights the antagonistic way in which the federal association of statutory sickness funds is viewed by the pharmaceutical stakeholders in Germany.

The frequency with which this issue was raised in the interviews justifies mentioning it as a possible auxiliary variable, but it is not without caveats. The main caveat is that the dominance of the federal association of statutory sickness funds was mainly raised as an issue by pharmaceutical manufacturers. It might therefore be a reflection of a phenomenon that one might usually expect in a corporatist system that relies heavily on bargaining structures and the power distributions these gives rise to. More research would have to be undertaken to determine exactly how the institutional distribution of bargaining power affects the outcome of pharmaceutical benefit assessments in Germany, but such research is beyond the scope of this thesis.

The second caveat is the fact that the criticism raised by pharmaceutical representatives in Germany fails to take account of the fact that, at the two crucial stages of the assessment process, i.e. the determination of the appropriate comparator and the appraisal of the added benefit, the federal association of statutory sickness funds is not the sole decision-maker in the process. That is to say, that it represents only one member group in the FJC, the other member groups are the physician and the hospital associations. While the federal association undoubtedly is a big player in making the decision on the choice of comparator and in deciding on the added benefit of a product, it is not the only player.

In terms of the values of accountability and transparency, the criticisms voiced by pharmaceutical manufacturers in relation to the question whether a member of the FJC should be involved in the price negotiations, appear legitimate. However, it is more difficult to understand why the federal association of statutory sickness funds should be conceived of as the sole culprit of pursuing one's own interests in the process. The fact that stakeholders raise such concerns about the dominant role of the federal association of statutory sickness funds demonstrates that there are underlying concerns amongst the stakeholder community about the institutional basis of the HTA paradigm in Germany. The questions of whether or not these concerns are justified and how the dominance of the payers' association in the benefit assessment process impacts on the outcome of assessment processes requires further in-depth research that is beyond the scope of this

thesis. If future research can confirm that the dominance of the federal association of statutory sickness funds indeed contributes to benefit assessment outcomes, this would be a significant finding in relation to the role played by institutional variables in Germany.

In addition to the perceived dominance of the federal association of statutory sickness funds in the early benefit assessment process, the empirical data suggests a dissonance between how the theoretical premise of ‘patient relevance’, the key idea within the German HTA policy paradigm, is perceived by stakeholders and how it is operationalised by decision-makers in practice. This dissonance arises from the debates around the definition and operationalisation of the concept of patient relevance in case-specific contexts and how a stringent and formal approach to defining the concept helps explain the outcome of pharmaceutical benefit assessments in certain cases.

As outlined in chapter 5, the added benefit of a pharmaceutical product is measured by its patient relevant therapeutic effect, which is expressed by way of patient relevant clinical endpoints. The main endpoints that the FJC and IQWiG consider patient relevant are mortality, morbidity and quality of life in terms of side effects. In each case of an early benefit assessment IQWiG and the FJC decide which clinical endpoints, in addition to the ones just mentioned, will be considered patient relevant. In addition to the centrality of mortality, morbidity and quality of life the two HTA bodies operationalise the concept of patient relevance by focusing on what the patient feels and whether this can be measured. According to a representative of the FJC this explains why, for example, laboratory parameters that are included in a given clinical trial will not automatically be considered patient relevant:

[...] laboratory parameters alone are not usually considered patient relevant. [...] we had big discussions about this in the case of hepatitis and the virus load [...] do I have hepatitis if I can detect it [in the blood] or not? [...] what is symptomatic, what the patient feels, quality of life etc. [...] that is patient relevant (Interviewee No. 16, 2013, p. 3).

The above quote helps explain why, in the case of Telaprevir and Boceprevir, IQWiG only accepted the main clinical endpoint of sustained virological response (SVR) as a surrogate endpoint for the patient relevant endpoint of hepatocellular carcinoma, but not as a patient relevant endpoint in itself. In IQWiG’s view, the clearance of the hepatitis C virus from the blood does not allow for an accurate

measurement of how the patient feels or whether he/she will develop follow-on complications from the hepatitis C infection in the future, which is why SVR in itself could not be considered patient relevant (IQWiG, 2012c). Similarly, IQWiG and the FJC do not usually accept the endpoint ‘progression-free survival’ as a patient relevant endpoint in the assessment of oncological products because it does not accurately express what a patient feels:

[...] what the patients do not feel [...] is not patient relevant. [...] things that patients do not feel but that only become apparent through diagnostic imaging, through laboratory diagnostics [...] we would say [...] progression-free survival in the case of oncological therapies in which the progression is only determined through diagnostic imaging [...] is not a patient relevant endpoint per se [...] (Interviewee No. 23, 2013, p. 5).

The above examples illustrate how patient relevance is conceptualised and operationalised by IQWiG and the FJC. However, this conceptualisation and operationalisation is criticised by the stakeholders as being too narrow in what is considered patient relevant. For example, the form of administration, i.e. intravenous or oral, of a pharmaceutical product is not considered patient relevant by IQWiG and the FJC because it cannot be measured as a clinical endpoint. This meant that in the case of Abiraterone and Fingolimod the form of administration was considered as an additional benefit of the treatment by NICE (2012d) while IQWiG (2012e) and the FJC (G-BA, 2012g) did not consider that this aspect gave rise to additional benefits. While the term ‘patient relevance’ led pharmaceutical manufacturers and other stakeholders to argue that the form of administration could be considered relevant to the patients (G-BA, 2012g), IQWiG and the FJC did not see its effect in any clinical endpoints and concluded that the effects of the form of administration on the actual state of health of the patients could not be measured.

Connected to the criticism that IQWiG’s and the FJC’s conceptualisation, and especially their operationalisation, of patient relevance is the criticism that the involvement of patients in the early benefit assessment is comparatively underdeveloped. Stakeholders in Germany question the legitimacy of the use of the concept of patient relevance if patient representatives are not asked about their view of what is patient relevant in specific cases. This seems to reflect an institutional situation in which patient groups do not make a wide use of the involvement mechanisms that are

available to them and in which an insitutionalised patient representation in the FJC speaks for all patients rather than for specific patients with certain diseases. The following excerpts of the stakeholder interviews exemplify how the above issues are being framed in the discussion of what determines pharmaceutical benefit outcomes in Germany:

[...] typically the industry fights with the FJC or IQWiG about the patient relevance. We say one thing, they say another, but de facto only patients can answer the question whether something is relevant for him (Interviewee No. 14, 2013, p. 4)

[...] we do not notice the patient representation [...], that may be a result of the degree of organisation considering the patient groups often have voluntary members who cannot get involved so much due to time constraints. [...] if you would give the patients a stronger voice in the FJC or on the topic of patient relevance, but I think one would have to formalise this. One cannot say we will wait until the patients get involved and then we will consider it somehow [...]. Instead [...] the impetus has to come from the institutional side to say, we want to consider patients more and that is why we are creating the structure that makes this possible (Interviewee No. 12, 2013, p. 8).

The brief overview of how the issue of patient relevance and the associated issue of patient involvement are framed in the German HTA context is relevant as an auxiliary variable to answering this thesis' research question because it suggests that the different ways in which the key concept of patient relevance is interpreted by the different stakeholders plays a role in what clinical endpoints are considered by the FJC as part of a benefit assessment. The inclusion or exclusion of clinical endpoints has a direct impact on how strong or weak the evidence base is, which in turn impacts on the outcome of a pharmaceutical benefit assessment. Empirically this means that IQWiG's and the FJC's operationalisation of patient relevance contributes to the final outcome of pharmaceutical benefit assessment. This is directly connected to how concepts and values are defined within the rules of evidence that were discussed in the previous chapters. As such the significance of patient relevance is an extension of the previously discussed rules of evidence. However, what is striking about concept of patient relevance is the heated debate amongst stakeholders about what *is* and what *should* be considered patient relevant, which gives rise to a within-paradigm dissonance that is described below.

The conceptualisation and operationalisation of patient relevance in the German system gives rise to a dissonance between the perceived theory and practice of the HTA paradigm. This observation is of great relevance to the theoretical premise of this thesis in that there seems to be confusion amongst stakeholders around the meaning and implications of the *idea* of patient relevance, which indicates a within-paradigm contestation of this term. This contestation gives rise to questions about the institutional basis of the Germany benefit assessment system and the comparative lack of involvement of patients.

Three scenarios are imaginable when it comes to resolving the apparent dissonance within the German HTA paradigm. The first one is a further specification as to how the FJC and IQWiG conceptualise and operationalise patient relevance, the second one is a pragmatic approach in which the two organisations gradually widen the operationalisation of the term and the third one is an institutional solution in which patients are given more of a voice on how patient relevance should be understood. An assessment of the likelihood of each of the scenarios is beyond the scope of this thesis. For the purpose of this thesis, the key finding that arises from the above discussion is that, in contrast to the English paradigm, the empirical evidence suggests that key ideas of the German HTA paradigm are contested in practice, which a) leads to a dissonance between paradigm theory and practice and b) might lead to a revision of key concepts in the German HTA paradigm to account for the current definitional and operational challenges in the future.

### **8.3.1. Research Findings**

The empirical evidence, and especially the interviews with stakeholders from the pharmaceutical industry in Germany, gives rise to the strength of bargaining powers of certain stakeholders vis a vis others as a contributor to the outcome of pharmaceutical benefit assessments. For the purposes of comprehensiveness this theme was included in the empirical section of this thesis. However, due to the fact that it was predominantly raised as an possible explanatory variable by pharmaceutical stakeholders, the true impact of the relative bargaining power of certain stakeholders vis a vis others in the bargaining process cannot be conclusively assessed within the scope of this thesis. Future research is needed to assess the impact of this issue and such research would need to include an analysis of documents such as minutes of FJC meetings in which the

appropriate comparator is discussed and minutes of negotiations between the sickness insurance funds and the pharmaceutical manufacturers. This data is currently not available in the public domain, which further exacerbates the problem of assessing the role that the bargaining power of certain stakeholders plays in determining the outcome of pharmaceutical benefit assessments in Germany.

Despite the aforementioned caveats in relation to ‘bargaining power’ as an auxiliary variable, the fact that it was raised by stakeholders in the interview process and in FJC hearings contributed to the research findings that are presented in this thesis. As previously discussed, one of this thesis’ research findings is that HTA policy paradigms are different in different countries and that this difference reflects institutional and ideational differences of the wider health care paradigm. Regardless of the fact that the issue of ‘bargaining power’ arose predominantly in interviews with pharmaceutical stakeholders, the fact that it arose at all reflects the institutional paradigmatic construct of the German system as a corporatist system. Corporatist health care systems are based on bargaining structures, powers and struggles and it is not surprising that at least one stakeholder group views the institutional structure as a contributing factor in pharmaceutical benefit assessments. Vice versa, the fact that this issue does not arise in the English context reflects a system in which struggles over bargaining power are less pronounced. Thus, the fact that the auxiliary variables that arise from the empirical evidence differ in Germany and England further supports the conclusion that HTA policy paradigms are different and that they reflect that wider health care paradigms.

The controversy around patient relevance acts as an auxiliary variable in Germany. This is mainly because, as was highlighted in the previous chapters, ‘patient relevance’ is a recurring sub-theme in the majority of the themes that arise from the empirical data. The above empirical examples in relation to the cases of Fingolimod, Abiraterone, Telaprevir and Boceprevir reaffirm the central role of the idea of patient relevance as an overriding value within the German HTA paradigm. However, in addition to this reaffirmation, the discussion presented in this section in relation to how the conceptualisation and operationalisation of patient relevance is contested gives rise to a further noteworthy research finding. In relation to the value, or idea of patient relevance, there is a dissonance between the theory and the practice of the HTA paradigm that manifests itself by different stakeholders conceptualising patient relevance differently, with some arguing that the concept should be understood more

literally to include patients' views more widely. However, the illustrative quotes by representatives from IQWiG and the FJC suggest that currently the principle of patient relevance should neither be equated with patient preferences nor with aspects of patient involvement. The principle should solely be viewed within the context of clinical benefits.

The exact effect of the described dissonance on the ultimate outcome of pharmaceutical benefit assessments is difficult to evaluate. However, the empirical evidence presented in this thesis suggests that the operationalisation of the concept of patient relevance has a big impact on which evidence is included in an assessment, that is to say that clinical trials are only included if the clinical endpoints are considered patient relevant. This in turn has a big impact on the outcome of the pharmaceutical benefit assessment as this is based on the available evidence.

More importantly though, the dissonance between theory and practice within the German HTA paradigm has significant theoretical implications. It suggests a disagreement within the stakeholder community on how the concept of patient relevance should be articulated and operationalised in normal decision-making. Except for general criteria such as mortality, morbidity and quality of life policy-makers in Germany did not specify how the concept of patient relevance could, or should, be operationalised. This led to an articulation of the paradigm in a way that does not take account of patient preferences and patient involvement. This in turn is criticised by stakeholders who question whether the operationalisation of patient relevance is in line with the wider paradigm. Thus, while the patient relevance paradigm can be described as dominant in the German context, multiple understandings of patient relevance appear to exist. This demonstrates that an idea such as patient relevance can mean different things to different stakeholders. It implies that paradigms can face within-paradigm contestation that might ultimately lead to changing or further developing certain principles. This can be expected for the principle of patient relevance in Germany.

In summary, the discussion of the auxiliary variables of 'bargaining power' and the conceptualisation of patient relevance in the German context gives rise to the following research findings:

1. The empirical evidence pertaining to pharmaceutical benefit assessment processes in Germany gives rise to two potential auxiliary variables: the relative bargaining power of certain stakeholders vis a vis others and the conceptualisation of patient relevance. While the validity of the first variable is

difficult to assess because of a potential bias arising from the fact that only pharmaceutical representatives raised it and because of a lack of available data to validate the claim, the contributory role played by second variable is less caveated because it arises, in different formats, throughout the empirical evidence that was analysed.

2. The contrast between auxiliary variables in England and in Germany suggests that the HTA policy paradigms are different in both countries and that this difference determines which auxiliary variables might contribute to the outcome of pharmaceutical benefit assessments. The difference also reflects the wider health care paradigms in both countries in that the auxiliary variables outlined in the case of German are logical extensions of a corporatist health care system, whereas the role of public pressure in England suggests that accountability towards the patients and the public dominates the sphere of auxiliary variables.
3. Finally, the importance of the definition and operationalisation of the idea of patient relevance in the German HTA paradigm has significant theoretical implications with regards to the theory and practice of policy paradigms. The German case highlights that within-paradigm dissonances and contestations can exist, meaning that paradigms are fluid rather than rigid constructs that may develop and change during normal decision-making processes, depending on how severe the dissonances between theory and practice are. This underlines Kuhn's (1962) point that a paradigm can only be understood by examining its application in normal science or, for the purpose of this thesis, in normal decision-making processes.

In the next and final chapter I draw together the conclusions and research findings that the previous empirical analysis gave rise to and outline areas for further research into what determines the outcomes of pharmaceutical benefit assessments in countries that employ formalised HTA procedures.

## Chapter 9

### Conclusion

#### 9.0. Introduction

The preceding chapters demonstrated that HTA paradigms in England and Germany are different (table 9.1.)<sup>36</sup>. In fact, except for the underlying rationale for conducting HTAs, the importance of thresholds (albeit in different formats) and a value core, table 9.1. suggests more differences than similarities when it comes to institutional and ideational features that characterise HTA paradigms. Based on the identification of the HTA paradigms one might therefore conclude that these differences play a role in determining the outcome of pharmaceutical benefit assessments. However, the six themes (table 6.1.) that emerged from the empirical analysis indicate more similarities in what matters with regards to the operationalisation of the paradigms than might have been expected, therefore underlining the theoretical premise that the mere identification of a paradigm is insufficient when it comes to understanding its effect. In practice, that is when paradigms are established and applied in normal decision-making, only three categories of factors (table 9.2.) appear to matter. These are the rules of evidence that paradigms give rise to, the core and periphery of values and procedural characteristics such as patient involvement and public pressure. This implies that paradigms are nuanced in the way they operate. By examining how they are applied in practice, we learn about their crucial features and what distinguishes them from those in other national contexts.

There is a striking similarity in the *kind* of issues that arose in the consultation and decision-making processes in the ten embedded cases. The broad themes that emerged were identical. Five out of six themes centred on questions of the quality, appropriateness and interpretation of evidence. However, the way the HTA agencies dealt with these themes and addressed the issues was different, thus suggesting that different rules of evidence arise in the process of articulating HTA paradigms. The

---

<sup>36</sup> Table 9.1. depicts the main features of the HTA paradigms in Germany and England as they were identified in chapters 4 and 5. They are presented in relation to the potential independent variables outlined in table 3.2. (chapter 3) in order to give an indication of the variables that would affect the outcome of pharmaceutical benefit assessments according to the HTA paradigm.

empirical analysis demonstrates that the outcome of pharmaceutical benefit assessments in health care systems that employ formalised HTA procedures is determined by how a similar set of themes around the evidence base on a pharmaceutical product gets transformed, interpreted and framed in the context of a given HTA paradigm.

In this final chapter I present and discuss the research findings that gave rise to the aforementioned conclusion. There are eight findings that emerged from the empirical analysis. The findings were interpreted with reference to the theoretical framework presented in chapter 2, the HTA paradigms outlined in chapters 4 and 5 and the dimensions of HTAs that were introduced in chapter 1. Following the discussion of the research findings, I provide an overview of the implications of the research, its limitations and possible areas for future study.

**TABLE 9.1. – HTA Paradigms England and Germany** (See chapters 4 and 5 for detailed discussion)

<b>Features of the HTA paradigm</b>	<b>England</b>	<b>Germany</b>
Decision-making in health care (=Institutional feature (IV2))	Centralised, national decision-making on the coverage of new pharmaceuticals. Localised decision-making on implementation of national guidelines and health care service provision.	Corporatist decision-making structures based on bargaining between providers and payers (statutory sickness insurance funds)
Challenges to the health care system, i.e. resource considerations (=Institutional feature (IV2))	Acknowledgement of finite health care resources in NHS Constitution	Lack of acknowledgement of finite health care resources in Social Code Book V
Cost containment policies (=Institutional feature (IV2))	Long history of cost containment and priority setting in health care. Expressed in terms of ‘value for money’ in NHS Constitution.	Priority setting not an established feature. Controversies around whether priority setting is occurring. Nikolaus-case gave rise to a right to health care treatment even if evidence of effectiveness is lacking.
Values & principles (=Ideational features (IV1))	Reference to wider social values such as equality and non-discrimination in NHS Constitution.	Reference to health-specific values such as right to treatment and solidarity.
Purpose of HTA (=Institutional features (IV2))	To recommend the in- or exclusion of a pharmaceutical product based on its clinical and cost effectiveness.	To inform price setting negotiations between the sickness insurance funds and the pharmaceutical manufacturers.
Procedural values of HTA (=Institutional features (IV2))	Importance of accountability and transparency recognised. Stakeholders involved in each of the numerous steps of HTA, i.e. in assessment and appraisal stages.	Importance of transparency recognised. Procedural steps of HTA limited. Stakeholder involvement limited to hearings in assessment stage, no involvement in appraisal stage.
Rationale of HTA (IV 2 & 3)	There can and should be a limit on what a health care state (Moran, 1999) should be willing to pay.	There can and should be a limit on what a health care state (Moran, 1999) should be willing to pay.
HTA criteria and principles (=Institutional and ideational features (IV 3))	Clinical and cost effectiveness. Thresholds in the form of ICERs. Thresholds conceptualised as thresholds for ‘value for money’	Clinical effectiveness conceptualised as the therapeutic patient relevant effect. Six categories of added benefit. IQWiG employs own method for operationalisation of thresholds between benefit categories. Thresholds conceptualised as thresholds for ‘right’ price.
Core and periphery of values (IV 1 & IV 3)	Yes, cost effectiveness as the most important value. Other values such innovation and social values can be considered.	Yes, patient relevance. Benefit categories express other values such as innovation.

**TABLE 9.2. – Paradigms in Normal HTA Decision-Making in England and Germany** (see chapters 6-8 for detailed discussion)

<b>Operationalisation and articulation of the paradigm based on case study analyses of what factors that determine the outcome of pharmaceutical benefit assessments.</b>	<b>England</b>	<b>Germany</b>
HTA criteria and principles (=IV 3)	The majority of factors that are considered are related to questions of evidence, the quality and the interpretation thereof. Rules of evidence are applied, i.e. use of thresholds in the form of ICERs. Rules of evidence are the expression of the HTA paradigm.	The majority of factors that are considered are related to questions of evidence, the quality of evidence and the interpretation thereof. Rules of evidence are applied, i.e. the need for the comparator product to reflect the licensed indication and the non-consideration of laboratory parameters as patient relevant endpoints. Rules of evidence are the expression of the HTA paradigm.
Core and periphery values (=IV 1 & IV 3)	Cost effectiveness (=core) has to be ensured. End-of-life criteria (=periphery) are applied in order to ensure cost effectiveness in some cases. Once cost effectiveness is given, other factors such as innovation (=periphery) can be considered.	Patient relevance (=core) has to be ensured. Controversies around conceptualisation and operationalisation of the term. Other principles do not play a large role. This suggests no significant values at the periphery of decision-making.
Procedural values (=IV 2)	Public pressure and stakeholder involvement may affect the outcome in exceptional cases.	Distribution of bargaining powers between stakeholders may affect the outcome in exceptional cases.

## **9.1. Research Findings**

The six themes (table 6.1.) that were identified in the empirical data can be subsumed under three features of the HTA paradigms (table 9.2.). During the process of normal decision-making, the factors that matter the most in determining the outcome of benefit assessments are the criteria of HTA in the form of rules of evidence, the core and periphery of values and, in exceptional cases, procedural values with regards to stakeholder involvement. All of these factors emerged as relevant in relation to themes that centred on questions of evidence. This indicates that the factors that play a role in normal decision-making are more limited than the principles that are contained in the overall paradigm. In HTA decision-making processes principles that require specification such as the question of how to operationalise thresholds play an important role, whereas other features of the HTA paradigm such as equality or solidarity appear to be adhered to implicitly but not referred to explicitly.

What follows is an overview of the eight research findings that emerged from this study, which will qualify the above statements in more detail.

### **9.1.1. Research Finding One: ‘Rules of Evidence’ Matter**

The empirical evidence suggests that the most important variable in determining the outcome of pharmaceutical benefit assessments is what Majone (1989) refers to as the ‘rules of evidence’. What constitutes evidence as part of a pharmaceutical benefit assessment, how it is defined and the criteria by which it is interpreted play a key role in determining the outcome of assessments. This finding is supported by the fact that five out of six themes that emerged from the empirical data centred on questions of evidence. The *kind* of issues that were discussed was similar in England and Germany. However, the way in which the issues were addressed differed. This difference can be explained by different rules of evidence that are applied in the process of normal decision-making.

The three empirical chapters discussed the six themes that emerged from the data. Chapter 6 outlined how evidence is conceptualised in different HTA policy paradigm contexts and how this impacts on the final outcome of pharmaceutical benefit assessments. Chapter 7 dealt with operational questions such as the definition of patient subgroups and comparator products for the purpose of assessments. Chapter 8 extended

the discussion by presenting auxiliary variables that might play a role in individual cases of pharmaceutical benefit assessments. Each of the chapters' conclusions highlighted the importance of the so-called rules of evidence for the final decision in a given case.

The importance of the rules of evidence is exemplified by how a given HTA policy paradigm responds to 'special' cases in which long-term effects of a given medicine cannot be captured accurately by the average length of randomised controlled trials (RCTs). Decisions on whether or not to use certain pieces of evidence ultimately have an effect on the final decision of pharmaceutical benefit assessments as the exclusion of certain pieces of evidence precludes the consideration of their data when making decisions. This is illustrated by the cases of Retigabine and Fingolimod, in which NICE and the FJC came to different decisions because the FJC did not consider all, or in Retigabine's case any, of the evidence submitted by the pharmaceutical manufacturer.

In relation to the theoretical framework the above remarks suggest that 'instrument settings' (Hall, 1993) have the most effect on the outcome of HTAs. This implies that the methodological choices made by policy-makers and decision-makers are key to understanding the outcome and effect of HTA processes.

### **9.1.2. Research Finding Two: HTA Paradigms Take Different Forms**

While the 'rules of evidence' matter in both Germany and England, the format they take differs. This demonstrates how HTA paradigms take different formats even when they are applied to broadly similar themes such as evidential questions. The difference can be explained by reference to the values and goals that are embedded in the wider policy paradigms of HTA and health care. For example, the use of ICERs by NICE reflects a health care paradigm in which 'value for money' has a long-standing tradition, whereas the FJC's use of benefit categories is a logical extension of a health care system that is based on bargaining between payers and providers. The FJC's benefit categories may appear unspecific, but they are a means of aiding the decision-making process without undermining the institutional bargaining structure that is inherent to the German health care system. In other words, the use of benefit categories ensures that the institutional features of the Germany paradigm remain intact.

Chapters 6, 7 and 8 provided a detailed discussion of the differences in the conceptualisation and operationalisation of the rules of evidence under a given HTA paradigm. The differences include how clinical effectiveness is conceptualised, how uncertainties in the evidence are dealt with and what pieces of evidence are considered in the absence of good-quality evidence. The discussion contained in these chapters gives rise to the conclusion that England's HTA paradigm is one of cost effectiveness, whereas Germany's is one of patient relevance. These are the dominant paradigms in England and Germany. This finding supports the branch of literature that highlights that different paradigms can exist in the same policy field (see chapter 2). However, the identification of national auxiliary paradigms would require more research than was feasible within the context of this thesis. Suffice it to say in this regard that the consultation documents in Germany suggest that the pharmaceutical industry may be advocating a paradigm in which cost savings are considered more than is currently the case.

### **9.1.3. Research Finding Three: A Similarity in Outcome Should not be Interpreted as a Similarity in Reasoning, or: Contrasting HTA Paradigms do not Necessarily Lead to Contrasting Outcomes**

The case of Telaprevir, and to a lesser extent the cases of Cabazitaxel, Eribulin and Ipilimumab, illustrates that different reasons, and different reasoning, in the decision-making processes does not preclude the possibility that the assessment outcomes in different countries may be similar. This suggests that there is more than one causal path to a positive, or negative, outcome in HTAs. It also underlines the importance of not sampling cases on the dependent variable, an approach that is tempting when examining HTAs as different cases are perceived as the most interesting cases. However, this thesis shows that researchers can learn valuable lessons from cases in which HTA outcomes are similar. The crucial lesson is that different rules of evidence do not, as a matter of course, lead to different outcomes. Moreover, the comparative outcomes of a HTA cannot be understood without an examination of the argumentative processes that led to it. In other words, one cannot make inferences about the independent variables based on the similarity or dissimilarity in the dependent variable. The empirical analysis presented in chapters 6-8 underscores that

pharmaceutical benefit assessments are the result of complex applications of paradigms rather than of one or two factors being present or absent in a specific case.

#### **9.1.4. Research Finding Four: Core and Periphery Values Exist**

Chapter 5 gave rise to the question whether there are principles and values in a paradigm that are weighted more in comparison to others, thus a core and periphery of values, or a minimum threshold of criteria, that have to be met. The empirical analysis supports the theory that core and periphery values exist. This becomes apparent when examining the thresholds or hurdles that evidence in a given HTA process needs to meet in order to receive a positive appraisal. The terms ‘core’ and ‘periphery’ of values are used here for want of better words to describe the observation that certain principles appear to be given more weight than others under the English and German HTA paradigms.

In England the principle that forms the core of decision-making is that of cost effectiveness. In Germany the most important principle is that of patient relevance. The empirical evidence suggests that it is more difficult for pharmaceutical products to receive a positive appraisal if the above values are not addressed in manner that satisfies the decision-making criteria. In fact, all of the positive appraisals included in this thesis were explained with reference to cost effectiveness or patient relevance respectively.

Whilst a core of values exists in England and Germany with regards to one principle that appears to carry more weight than others, a difference is apparent when it comes to periphery values. Whereas England’s decision-making paradigm does in fact appear to allow for the consideration of principles such as innovation and social values as periphery decision-making values once cost effectiveness is met, the empirical evidence in Germany does not suggest that other principles are considered once the criterion patient relevance is satisfied. This appears to be connected to the conceptualisation of the benefit categories as incorporating the most important outcomes that the treatment with a new medicine can have, thus implying no need for additional, or periphery, values to be considered outside the core benefit categories.

Finally, both the theory and the practice of the English paradigm include contingencies for when the usual cost effectiveness threshold is not met. Values such as innovation and, most frequently, end-of-life considerations are invoked if cost effectiveness is not met. This allows for flexibility in operationalising the paradigm in a

way that alters the threshold criteria by reference to other values in cases that warrant it. In contrast, the empirical evidence in Germany suggests that invoking principles other than patient relevance is a difficult to impossible endeavour. From a theoretical point of view this is an interesting finding as it reveals that, at least in the case of HTAs, paradigms are further articulated in normal decision-making when uncertainties arise or when thresholds are not met. This is in line with Kuhn's (1962) thoughts on the importance of 'normal science' as a process in which tools, criteria and concepts of a paradigm are refined. As a result, NICE has developed its rules of evidence to allow for the consideration of other values to supplement cost effectiveness considerations, whilst the FJC and IQWiG are sticking to the concept of patient relevance as the main paradigmatic principle.

However, the point on the flexibility, or lack thereof, of paradigms should not be overemphasised as it may be a result of the fact that NICE has existed longer and has therefore had more opportunities to further articulate its paradigm. In time, the FJC and IQWiG may (have to) do the same. Alternatively, if it does not, it would make for an interesting study on how certain institutional, cultural or societal prohibit it from further articulating its HTA paradigm in a way that addresses potential shortcomings of given pieces of evidence.

#### **9.1.5. Research Finding Five: Thresholds as the Expression of Paradigms in Normal Decision-Making**

The importance of thresholds for HTA paradigms cannot be overemphasised. Against the background of the underlying rationale for HTAs, i.e. HTAs are about finding out whether the use of a product promises *enough* clinical benefit to justify a certain price or the inclusion in a health care benefit basket, thresholds are the means by which the limits of clinical benefit or cost effectiveness are given meaning. That is not to say that thresholds are uncontroversial, indeed Appleby et al. (2009) point out that the method, or lack thereof, by which they are set is frequently criticised. However, the point here is not to justify whether certain threshold levels are appropriate or not but to highlight their importance as an expression of the dominant HTA paradigms. In each and every case that was analysed in this thesis, the products had to meet the thresholds of cost effectiveness and patient relevance respectively. The HTA paradigms that were identified derive their labels from those thresholds as they were the ultimate hurdles the

evidence had to surpass. This finding calls for further research that includes a larger number of cases in order to examine whether it can be generalised, i.e. whether HTA paradigms always take the form of the thresholds they contain.

#### **9.1.6. Research Finding Six: Paradigmatic Coherence Matters**

While both the cost effectiveness paradigm and the patient relevance paradigms arise from the ideational and institutional characteristics of their respective health care systems, in Germany there appears to be a dissonance between the theory and practice of how the paradigm operates. As outlined in chapter 5, the criterion that health care decisions are to be based on the ‘generally accepted state of medical knowledge’ is contained in Social Code Book V (BMJV, 2013). For the purpose of HTA, this criterion is operationalised by using the principle of patient relevance. However, the way it is operationalised is frequently perceived as a mismatch between theory and practice. Stakeholders criticise the way patient relevance is operationalised as giving inadequate recognition of the ‘generally accepted state of medical knowledge’. For example in the case of Telaprevir the main clinical endpoint was only accepted as a surrogate parameter rather than an endpoint in its own right (chapter 6 and 7) even though clinical experts underlined that it is generally accepted by the international medical community as an endpoint in its own right (G-BA, 2012a). The perceived mismatch between theory and practice in the Germany paradigm gives rise to controversies and insecurities that are yet to be resolved.

The way thresholds and HTA criteria are operationalised impacts on the outcome of pharmaceutical benefit assessment. However, the finding described in the previous paragraph has theoretical implications as it demonstrates that paradigms are not static constructs. Instead they require refinement and articulation that may or may not give rise to alterations to rules of evidence. Moreover, it highlights the importance of paradigms being coherent within themselves, an importance that has previously been raised by Hall (1993, p. 290).

### **9.1.7. Research Finding Seven: Ideational and Institutional Variables Matter**

As the overview in table 9.2. shows the themes that emerged in the empirical analysis can be characterised as reflecting both ideational and institutional variables. Ideational variables matter to the outcome of benefit assessments in that the rules of evidence reflect a value judgement about considering certain principles and not others. For example, the use of ICERs in England reflects a value judgement that ‘value for money’ has to be safeguarded whilst the use of benefit categories in Germany reflects a value judgement about comparing clinical benefits of medicines within diseases categories and not across them. Essentially, these choices reflect views on how one sees the world and on what the world around you, i.e. in its institutional constructs, allows you to consider.

The empirical evidence does not suggest a dominant role of institutional variables in the pharmaceutical benefit assessment process. However, this demands a word of caution as the role of institutional variables may be more implicit in nature. For example, the above mentioned choice for ICERs and patient relevance respectively reflects institutional goals of HTAs in England in Germany. Since the purpose of HTAs in England is to inform the decision on the in- or exclusion of a pharmaceutical product in the health care benefit basket, the use of QALYs and ICERs to compare cost effectiveness across disease categories gives a clear answer on what is considered a cost effective use of NHS resources. By contrast, the impact of the categorisation of clinical benefits in Germany is less clear. The FJC does not attach price recommendations to a benefit category, which means that it is still up to the sickness funds and the pharmaceutical manufacturers to negotiate an adequate price. In this way the methodological choices in England and Germany also reflect the institutional features of the respective paradigms.

Finally, the empirical discussion in chapter 8 gives rise to the question whether institutional variables such as public pressure and stakeholder involvement are more likely to matter in exceptional cases. In ‘normal’, or comparatively uncontroversial, cases the impact of public pressure and stakeholder involvement on the outcome of pharmaceutical benefit assessments is low. It appears to matter more in highly salient cases such as the case of Abiraterone in England. This observation gives rise to numerous complexities that need to be considered in future research. These complexities include the question whether the resources of the advocacy groups in the

case of Abiraterone were more abundant than in the other cases. However, it is noteworthy that the same advocacy groups embarked on an extensive media campaign in the case of Abiraterone, but not in the case of Cabazitaxel. In the context of this thesis, the role of public pressure, stakeholder involvement and the distribution of bargaining powers as an independent variable cannot be conclusively assessed as it only emerged as an auxiliary variable in a few cases.

#### **9.1.8. Research Finding Eight: Paradigms can Explain Empirical Phenomena other than Policy Change**

The use of paradigms to guide the interpretation of the empirical findings proved helpful in explaining the outcomes of pharmaceutical benefit assessments. The broad categorisation of variables into ideational and institutional features based on Hall's (1993) conceptualisation of policy paradigms helped maintain the construct for the analysis throughout the thesis. However, the effect of paradigms only really gained meaning in relation to this thesis' research question through an examination of normal decision-making processes as advocated by Kuhn (1962).

The empirical analysis of 'normal' decision-making processes in the form of pharmaceutical benefit assessments showed that paradigms function in a more nuanced and less static manner than is portrayed in the extant literature. Paradigms, and the rules of evidence therein, are articulated and redefined in HTA processes, especially when uncertainties that emerge from the evidence need to be addressed. In practice, the paradigmatic factors that determine the final outcome of pharmaceutical benefit assessments are less numerous than the mere identification of a paradigm would suggest. In practice, what matters are the rules of evidence, the core values and, to a lesser extent, the public pressure that is exerted. This conclusion was only possible by tracing the argumentative and reasoning processes in the ten embedded case studies.

The aforementioned assertions have implications for the branch of public policy literature that examines the role of paradigms. They suggest that a more in-depth understanding of paradigms can be gained by engaging in a three-step analytical process focused on identification, operationalisation and impact. That is to say that in addition to identifying a paradigm, its operationalisation has to be examined before evaluating its impact. Currently, most authors in the field focus on the identification and the impact without providing an account of what it is about a paradigm that matters in actual

decision-making. What I suggest makes the use of paradigms to explain empirical puzzles both more lengthy and more complex. However, in the case of this thesis, the approach gave rise to a specific and nuanced set of paradigmatic variables that determine the outcome of pharmaceutical benefit assessments. This provides a good basis for exploring whether a similar delineation of paradigm characteristics is possible in other policy areas.

## **9.2. Generalisability and Limitations of the Research Findings**

As with most qualitative studies, the generalisability of the research findings to other cases of pharmaceutical benefit assessments has to be viewed with caution because generalisability increases with a larger number of cases. The aim of this research was not the establishment of definite cause-and-effect relationships that would be generalisable to a wider set of cases, but the detailed analysis of the decision-making processes of benefit assessments to understand what determines the final outcome. Given that ‘evidence’ is at the core of HTA processes it is probable that the analysis of more cases of pharmaceutical benefit assessments would lead to similar findings as the ones presented here. However, further research is needed to determine whether the conclusion that the outcome of benefit assessments is determined by how a similar set of questions is addressed by different rules of evidence under given paradigms is valid for other cases. As outlined in research finding eight, such research would not only need to identify a given paradigm, but also examine how it is operationalised in order to understand its effect.

Despite the cautious remarks on generalisability, this thesis contains at least one significant finding, which is likely to be generalisable to other cases of HTAs. This is the finding (see 9.1.3.) that the score of the dependent variable does not necessarily allow for an account of the reasons that contributed to the outcome. As the case of Telaprevir demonstrated, the outcome in the dependent variable can be similar even if different rules of evidence are applied in the process. This has methodological implications in terms of generalisability. Inferences based on the dependent variable should be avoided in the field of pharmaceutical benefit assessments if one wishes to understand what shapes the outcomes.

There are three main limitations of the research findings. The first is theoretical in kind in that the question arises whether paradigms matter as much in decision-making

areas that are not strongly evidence-based. That is to say, are evidence-based policies more likely to be shaped by paradigms than other, perhaps more political and strategic, decision-making areas such as foreign or military policy? More research is required to address whether this is indeed a limitation of the research findings.

The second limitation relates to the finding that political factors are not a major contributor in determining the outcome of benefit assessments. The finding may be a result of under-representation or social desirability bias in the data. That is to say that decision-makers are unlikely to admit that a decision was made due to political concerns as this would undermine their organisation's standing as an independent, quasi-scientific decision-making body. This is connected to a third limitation, that is the lack of data to provide conclusive answers to some of the questions raised. For example, in Germany the proceedings of the committees that prepare the benefit assessment decisions for the FJC are made public only in summary form, meaning that the arguments brought forward in this process are not easily traceable. Therefore, the research findings might be limited due to social desirability bias and data restrictions. However, such limitations are not unique to this research. Overall, the research findings still provide a valuable contribution to the question of what determines the outcomes of pharmaceutical benefit assessments and how paradigms shape them. They also give rise to questions that future research should address, some of which have been alluded to in this section and more of which are outlined in the following section.

### **9.3. Future Areas of Research**

In addition to the topics outlined in the previous section, there are three key topics that demand further research. The first is the question that so-called 'issue characteristics' (Lowi, 1964) of medicines and the diseases they are indicated for play. Are value judgements and considerations different when the medicine being assessed is one for cancer versus one for, for example, blood pressure? This question demands further exploration because cancer treatments receive much public and financial attention. Much of the current research funds are dedicated to cancer research and debates around treatment, especially end of life treatment, for cancer patients are very emotional in the media. The majority of medicines undergoing HTAs are cancer treatments. This might be a compounding factor determining the outcome of benefit assessments on cancer treatments. The question emerges whether the research findings

contained in this thesis are generalisable to cases other than cancer treatments. As outlined in chapter 3, the sample of cases that was analysed was evenly spread in terms of the diseases for which the new interventions were indicated. Nevertheless, the fact that a large number of medicines that undergo benefit assessments are cancer drugs warrants an analysis on whether decision-making is somehow different in cases of different disease indications.

In this study that above question emerged in relation to the benefit assessments of the cancer interventions. The two cases (Cabazitaxel and Eribulin) that did not receive a positive appraisal from either NICE or the FJC were cancer treatments, more specifically chemotherapy treatments. Meanwhile, the two cancer treatments that did receive a positive appraisal were Abiraterone, which is a hormonal therapy, and Ipilimumab, which works by activating the immune system to fight cancer cells. Both of these treatments are very different from chemotherapies, which are associated with substantial side effects. The role that these issue characteristics play in the likelihood of whether a new medicine is recommended demands further exploration.

The second topic that demands further research arises from the auxiliary variables discussed in chapter 8. The role of these auxiliary variables would justify a research study in its own right and would most likely include issues of stakeholder and lobbying influence. However, at least in Germany's case, access to appropriate data would be challenging because the FJC's decision-making minutes are not publicly accessible and the price negotiations between sickness funds and pharmaceutical manufacturers are confidential. Nevertheless, it is a study waiting to be done. It is also a study that could employ theoretical frameworks such as Sabatier's advocacy coalition frameworks in order to examine whether the dominant HTA paradigms are more or less aligned with the belief systems of a given advocacy coalition.

The third area that demands further research is connected to the theoretical contribution that this thesis makes. The thesis demonstrates that ideational accounts of policy processes can help explain the outcome of complex decision-making processes and not just the outcome of processes of policy change. The relevance of this finding to other policy areas such as environment, defense or education policy is worth exploring in future research. If ideational accounts can explain decision-making processes in addition to change in different policy areas, this gives rise to a wide range of subject areas that could be explored from a new angle. This would demand a move away from

the preoccupation with the outcomes of dependent variables to an analysis of the causal *mechanisms* determine the outcomes of decisions.

#### **9.4. Policy Implications**

The main finding of this study concludes that questions around evidence, and how they are addressed by applying different rules of evidence, determines the outcome of pharmaceutical benefit assessments. In the proceeding paragraphs I distinguish between the policy implications this has for different stakeholders. The implications are different for different stakeholders in the kind of opportunities and challenges they present. While the policy implications for policy-makers are complex, they are positive and full of potential for the pharmaceutical industry and challenging for patient and public organisations.

The implications for policy-makers relate to the wider political goals and hopes that have been attached to the introduction of HTA policies. These goals include a) the hope that difficult health care priority setting decisions can be de-politicised by establishing independent scientific institutions to carry out evaluations of the best available evidence and b) that these evaluations will lead to improved decision-making and, potentially, efficiency savings by distinguishing between effective and ineffective treatments (see chapter 1). The empirical results suggest that the goal of ‘de-politicising’ decisions in health care has been achieved in both England and Germany in the sense that the outcome of pharmaceutical benefit assessments is determined by the rules of evidence that are applied under a given HTA paradigm. However, there are political developments in both countries that suggest that the effect of achieving this goal might not be as well perceived by policy-makers as was intended.

In 2010 the Conservative-Liberal Democratic Government in England created a Cancer Drugs Fund (CDF) of £200 million annually to provide patients access to cancer drugs that have not been recommended by NICE and that their clinician thinks they will benefit from (NHS England, 2014). Additionally, the Government intended to “[...] reform NICE and move to a system of value-based pricing, so that all patients can access the drugs and treatments their doctors think they need” (Cabinet Office, 2010, p. 25). Whilst the future of value-based pricing (VBP) – a pharmaceutical pricing system in which the price of a new medicine is to reflect its value – remains uncertain, both the

creation of a CDF and the move towards a system of VBP give rise to the question whether there is a dissatisfaction with the current HTA paradigm in England.

Politically the establishment of a CDF undermines NICE's paradigm in that it suggests that NICE's recommendations are, in some instances, not good enough. A policy observer summarises this in the following way:

[...] the fund had the political benefit of defusing the damaging arguments that have arisen when officials have denied patients access to expensive cancer treatment [...] on cost-benefit grounds (Cook, 2014).

Notwithstanding political opportunism as a motivating factor in the creation of the CDF, the fund represents a budget that is external to the rest of the NHS budget. It also represents a political decision to prioritise cancer patients over other patients such as patients with rare and/or chronic long-term conditions who might also benefit from a stand-alone fund that ensures access to medicines that are not considered cost effective by NICE. The latter raises the question of how the CDF is justified in relation to values such as equality and non-discrimination.

This study demonstrated that the outcomes of pharmaceutical benefit assessments are determined by methodological processes and choices. Political considerations and influence did not feature prominently in the data. However, the outlined developments in relation to the CDF and VBP imply that there is a trend towards re-politicising health priority setting because the HTA paradigm does not address the questions that policy-makers consider politically and ethically salient. This has wider policy implications in relation to the question of whether the de-politicisation of a policy area is ever truly possible. Authors such as Landwehr (2009) and Holm (1998) have argued that it is not possible in the field of HTA. The findings contained in this thesis suggest that although a de-politicisation is possible, in reality it may not be desirable. The described developments appear to indicate that the effect of de-politicising decisions may not be in line with political commitments that policy-makers wish to uphold. The implications of this are wide-ranging in that it questions the very reason for conducting HTAs in the first place. If policy-makers are happy to mitigate the effects of benefit assessment outcomes on the availability of medicines for certain patient populations by introducing policies to circumvent them, the question arises

whether financial and political resources should be spent on conducting HTAs in the first place.

Meanwhile, a political debate on the effects of early benefit assessments is developing in Germany, especially in relation to the implications of a new medicine being categorised in one of the higher benefit categories. As outlined in chapters 4 and 5, the outcomes of benefit assessments in Germany inform the price negotiations between the sickness funds and the pharmaceutical manufacturers. The system is based on the idea that the higher the benefit category, the higher the price that a pharmaceutical manufacturer can demand for the product in question. However, the political question that has recently emerged is whether a high benefit category automatically justifies the product being priced very highly, or, whether there can or should be limits on what a manufacturer can demand even if a product is considered highly innovative.

The above debate specifically arose around a new pharmaceutical product called Sofosbuvir (brand name Sovaldi), indicated for the treatment of chronic Hepatitis-C infections. The FJC assigned the second highest benefit category to Sofosbuvir (G-BA, 2014). Since this decision in February 2014, the pharmaceutical manufacturer of Sofosbuvir has been criticised for the high price it is demanding at €60.000 for a 12-week treatment course (Deutsche Apotheker Zeitung, 2014). Clinicians, statutory sickness funds and politicians are criticising the price as unacceptable. According to Jens Spahn, an expert on health policy in the current Government: “Sovaldi is a real innovation and a blessing for many patients, but that does not justify an astronomical price by any stretch of the imagination”<sup>37</sup> (Deutsche Apotheker Zeitung, 2014a). Similarly, NICE criticised the pharmaceutical manufacturer of the new breast cancer drug Kadcyla for not offering “[...] its new treatment at a price that would enable it to be available for routine use in the NHS” (NICE, 2014d). There appears to be a similar political debate about the price setting for new pharmaceuticals in Germany and in England.

In light of the research findings the outlined political debate is significant because it suggests that the current HTA paradigms do not adequately address this important factor. The pharmaceutical benefit assessment outcomes are determined by methodological and scientific questions in relation to the evidence on a product. This

---

<sup>37</sup> Translation provided by the author of this thesis.

implies that the issues that are considered are comparatively limited and do not address wider concerns such as the ‘right’ price for a product. Even in the German HTA system, where the benefit category is supposed to indicate whether a higher or lower price would be justified, the higher categories are not matched by a *category of prices*, which would provide guidance for the negotiating parties. In connection to pharmaceutical price setting policies, the thesis’ findings indicate that there is little room, or little willingness by decision-makers, to consider pharmaceutical prices under the current HTA paradigms in England and Germany. This implies that the situation can only be rectified by altering the current systems in a way that allows for the incorporation of price considerations within the rules of evidence. This is likely to only be possible if the political will is strong enough, thus underlining the likelihood of a re-politicisation of the area.

The final two policy implications that I draw attention to relate to the impact that dominant HTA paradigms of cost effectiveness and patient relevance have on stakeholders in the area. The existence of diverging HTA paradigms is challenging for pharmaceutical manufacturers because the evidence base on a given product has to be tailored to specific national paradigms. This can consume resources and time and is therefore frequently criticised as a negative outcome of the introduction of HTA policies. However, while there is no question that the adjustment to different HTA system requires time and resources, I argue that it ultimately presents opportunities to make evidence better and stronger. That is to say, that while the onus of adjusting the evidence base to new criteria lies with pharmaceutical manufacturers, they can contribute to better health care by considering a wider set of criteria and principles in clinical trials than is currently the case. The findings of this research imply that pharmaceutical manufacturers are in a good position to meet the demands of different HTA systems because they are the providers of evidence. The outcomes of pharmaceutical benefit assessments in England and Germany are determined by evidential questions, thus, while time-consuming, the HTA paradigms do not appear to constitute the huge black box of randomness and antagonism that the pharmaceutical industry occasionally still portrays them as.

Whilst the implications of the findings are potentially positive for the pharmaceutical industry, the implications are more challenging for stakeholders such as patient and public organisations. The finding that the outcome of pharmaceutical benefit assessments is based on complex and detailed interpretations of an evidence base gives

rise to the inference that patient and public views will carry more weight if they contribute to these interpretations. This in turn demands expertise, resources and skills from patient groups, the attainment of which might not be realistic or feasible. According to one interviewee in England:

[...] many if not most patient organisations [...] if they are of a certain size and can afford to employ people like me to contribute to something which is quite technical and quite specific [...] we do [...] address NICE in terms of [...] what evidence we provided, how we can interpret their appraisal [...] patient organisations have themselves become experts in NICE processes (Interviewee No. 19, 2013, p. 4).

There is no guarantee that every patient group will be able to address the issues that a HTA body considers important in the way that is described in the above interview. For example, an interviewee in Germany asserted:

[...] in Germany we [...] have the problem that patient organisations are not very well organised [...] there are a lot that are not in a position to see through the complex process, that depends on the financial and human resources that one has [...] and these people then come together with people who do this for a living [...] (Interviewee No. 8, 2013, p. 6).

The above remarks imply that the kind of patient and public participation and consultation that ultimately affects the outcome of benefit assessments is hard to achieve. While some patient organisations in England have risen to the challenge and have become experts themselves, this development is quite unique at this point. Politically, this calls for an examination of the status quo and of the true involvement that patient and public organisations have in the process of HTA. The question arises whether current HTA paradigms require alterations to better accommodate the more societal and ethical perspectives that patients contribute to the process.

## 9.5. Concluding Remarks

The eight research findings give rise to the conclusion that the outcome of pharmaceutical benefit assessments in health care systems that employ formalised HTA procedures is determined by how a similar set of themes around the evidence base on a pharmaceutical product gets transformed, interpreted and framed in the context of a given HTA paradigm. This conclusion contributes to the empirical knowledge on HTA procedures by underlining that evidence gains relevance and meaning within national contexts and does not carry relevance in its own right. Pharmaceutical companies, national governments, HTA decision-making bodies, clinicians and patient groups interpret the available evidence from their own unique perspectives. However, the dominant perspectives of how to interpret evidence are found in the dominant HTA paradigms in countries that employ HTA processes to inform decision-making in health care. The dominant HTA paradigms in England and Germany are the cost effectiveness and patient relevance paradigms respectively. The identification of these paradigms is another empirical contribution of this thesis.

In addition to contributing to a better empirical understanding of what determines the outcome of pharmaceutical benefit assessments, this thesis' most important contribution is theoretical. By using policy paradigm frameworks and Kuhn's (1962) concept of 'normal science' to analyse and interpret the results of this study, I demonstrated that theories of paradigms can help explain empirical phenomena other than policy change. Moreover, I demonstrated that the features and functions of paradigms can be understood better by analysing how they are operationalised in normal practice.

The themes that emerged from the empirical analysis showed they were limited to three sets of variables that were identified as part of the broader HTA paradigms (table 9.2.). In practice, what matters in determining the outcome of pharmaceutical benefit assessments are firstly, the criteria, or rules of evidence, in HTA decision-making, secondly, the values core within a HTA paradigm and thirdly, albeit to a lesser extent, procedural factors such as the involvement and pressure of stakeholders in the assessment process. This suggests that paradigms, when articulated and applied in normal decision-making processes, are more manageable, measurable and fluid intellectual constructs than the extant literature depicts them to be. It also suggests it is possible to describe the important independent variables of paradigms by examining

how they are operationalised. Future studies on policy paradigms would benefit from the analysis of how paradigms operate in ordinary, i.e. 'normal', processes and not just in the extraordinary processes of policy change.

## APPENDIX A

### Overview of NICE and Federal Joint Committee Decisions

Product & Indication	NICE	Federal Joint Committee
Abiraterone (Prostate Cancer)	<ul style="list-style-type: none"> <li>- Initially not recommended (ICER too high, End of life criteria not met)</li> <li>- Then recommended upon presentation of new evidence</li> <li>- End of life criteria met</li> <li>- Patient Access Scheme</li> <li>- Oral drug benefit not captured in QALY</li> <li>- ICER likely less than £50,000</li> </ul>	<ul style="list-style-type: none"> <li>- Divided up the patient population into two subpopulations</li> <li>- For ‘best supportive care’ population: Indication for a significant additional benefit</li> <li>- Docetaxel-re-therapy population: Additional benefit not substantiated, missing data</li> </ul>
Apixaban (Prevention of thromboembolic events after hip or knee replacements)	<ul style="list-style-type: none"> <li>- Recommended</li> <li>- More clinically effective and cheaper than at least one comparator (enoxaparin)</li> <li>- ICER not clear</li> </ul>	<ul style="list-style-type: none"> <li>- For knee-replacement population: No additional benefit</li> <li>- For hip replacement population: Marginal additional benefit</li> </ul>
Boceprevir (Chronic Hepatitis C genotype 1)	<ul style="list-style-type: none"> <li>- Recommended</li> <li>- ICERs all below £20,000/QALY</li> <li>- Clinically more effective than the comparator alone at achieving sustained virological response (SVR, taken as equivalent to a cure)</li> </ul>	<ul style="list-style-type: none"> <li>- For patient-naïve patients: additional benefit but not quantifiable</li> <li>- For therapy-experienced patients: additional benefit but not quantifiable</li> </ul>
Cabazitaxel (Prostate Cancer)	<ul style="list-style-type: none"> <li>- Not recommended, ICER too high at £87,500 QALY</li> <li>- Uncertainty about robustness of ICER</li> <li>- Effective, life-extending treatment but too much additional weight would have to be put on QALYs to make it an appropriate use of NHS resources</li> </ul>	<ul style="list-style-type: none"> <li>- Best-supportive care population: Indication for marginal additional benefit</li> <li>- Docetaxel-re-therapy population: No additional benefit substantiated, proof/data missing</li> </ul>
Eribulin (Advanced Breast)	<ul style="list-style-type: none"> <li>- Not recommended, ICER of £68,600/QALY likely</li> </ul>	<ul style="list-style-type: none"> <li>- Marginal additional benefit</li> </ul>

Cancer)	<ul style="list-style-type: none"> <li>underestimated</li> <li>- Concerns over toxicity profile</li> <li>- However, first drug in the population to show an overall survival benefit in an RCT</li> </ul>	
Fingolimod (Multiple Sclerosis)	<ul style="list-style-type: none"> <li>- Committee made an exceptional case for this drug, recommended it (after initially not recommending it and facing considerable protest)</li> <li>- Patient Access Scheme</li> <li>- Valuable new therapy, oral formulation is an innovation</li> <li>- £25,000-£35,000/QALY</li> </ul>	<ul style="list-style-type: none"> <li>- Marginal benefit only for patients with rapidly progressing severe MS, not for other patients such as patients with highly active remitting MS</li> </ul>
Ipilimumab (Advanced Melanoma)	<ul style="list-style-type: none"> <li>- Recommended</li> <li>- End of life criteria met</li> <li>- Patient Access Scheme</li> <li>- £42,200/QALY</li> <li>- Innovation because of few advances in advanced melanoma; currently the treatment option for these patients is enrolment in clinical trials</li> </ul>	<ul style="list-style-type: none"> <li>- Significant additional benefit</li> </ul>
Retigabin (Epilepsy)	<ul style="list-style-type: none"> <li>- Recommended</li> <li>- ICERs highly uncertain but novel mode of action and provision of new treatment option where others have failed</li> </ul>	<ul style="list-style-type: none"> <li>- Additional benefit not substantiated, missing data/proof</li> </ul>
Telaprevir (Chronic hepatitis C, genotype 1)	<ul style="list-style-type: none"> <li>- Recommended</li> <li>- ICERs low</li> <li>- Public health benefit highlighted</li> <li>- Stigma reduction etc.</li> </ul>	<ul style="list-style-type: none"> <li>- Additional benefit, but not quantifiable</li> </ul>
Ticagrelor (Acute coronary syndromes)	<ul style="list-style-type: none"> <li>- Recommended for all patient populations under consideration (4 in total)</li> <li>- All ICERs under £10,000/QALY</li> </ul>	<ul style="list-style-type: none"> <li>- Significant additional benefit for patients with non-ST-segment elevated myocardial infarction</li> <li>- No additional benefit in the other three patient populations</li> </ul>

## **APPENDIX B**

### **Interview protocol**

#### Open-ended questions for semi-structured interviews in England:

1. How would you describe your involvement [or the involvement of the institution/group you represent] in the assessment process of [insert name of pharmaceutical product]?
2. What were the main points you brought forward in the above process?
3. Do you feel your involvement in the process made a difference? [If yes, follow up on how it made a difference]
4. What determined NICE's ultimate decision in this process?
5. How would you describe the main principles that guide decision-making at NICE?
6. How would you describe the influence of other stakeholders in the assessment process? [Follow-up question: What is your relationship with these stakeholders?]
7. Do you feel that NICE does a good job at balancing the issues of clinical benefits and need, costs and innovation?
8. What role should NICE play when it comes to disinvestment incentives?
9. How do you view the future of health technology assessment in England, especially in light of the Government's plans to introduce a system of value-based pricing for pharmaceuticals?

#### Open-ended questions for semi-structured interviews in Germany:

1. How would you describe your involvement [or the involvement of the institution/group you represent] in the assessment process of [insert name of pharmaceutical product]?
2. What were the main points you brought forward in the above process?
3. Do you feel your involvement in the process made a difference? [If yes, follow up on how it made a difference]
4. What determined the FJC's ultimate decision in the process?
5. How would you describe the main principles that guide decision-making at IQWiG and the FJC?

6. How would you describe the influence of other stakeholders in the assessment process? [Follow-up questions: What is your relationship with these stakeholders? How would you describe the influence of patient organisations in the assessment process?]
7. What role does the requirement for ‘patient relevant’ outcomes play during the assessment process? [Follow-up: How would you define patient relevance?]
8. In your view, what role does the consideration of costs play in determining the outcome of pharmaceutical assessment processes?
9. What role have the early benefit assessments played in the price negotiations between sickness funds and pharmaceutical manufacturers?
10. How do you view the future of early benefit assessments of pharmaceuticals in Germany?

Themes explored with the above questions:

- Perception of guiding decision-making principles (generally and in the specific cases)
- Perception of stakeholders’ own involvement and influence in the process
- Perception of other stakeholders’ influence
- Perception of role of specific (paradigmatic) issues such as patient relevance, costs, balance between different considerations, methods
- Wider impact of HTA on pharmaceutical policy and health care system

## References

- Aaron, H.J. and Schwartz, W.B., 1984. *The Painful Prescription: Rationing Hospital Care*. Washington D.C.: The Brookings Institution.
- Abel-Smith, B., 1994. *Introduction to Health: Policy, Planning and Financing*. Harlow: Addison Wesley Longman Limited.
- Abraham, J., 2009. Partial Progress: Governing the Pharmaceutical Industry and the NHS, 1948-2008. *Journal of Health Politics, Policy and Law*, 34(6), pp. 931-977.
- Ärzteblatt, 2011. *Bahr erteilt Medizin nach Ranglisten klare Absage*. [online] Available at: <http://www.aerzteblatt.de/nachrichten/46052> [Accessed 26 January 2012].
- Altenstetter, C. and Busse, R., 2005. Health Care Reform in Germany: Patchwork Change within Established Structures. *Journal of Health Politics, Policy and Law*, 30(1-2), pp. 121-142.
- Anglo-German Foundation, 2008. *Publications*. [online] Available at: <http://www.agf.org.uk/publications/reports/2010/index.html> [Accessed 26 October 2014].
- Appleby, J., Devlin, N., Parkin, D., Buxton, M. and Chalkidou, K., 2009. Searching for cost effectiveness thresholds in the NHS. *Health Policy*, 91(3), pp. 239-245.
- BBC, 2014. *Breast cancer drug price reduction urged*. [online] Available at: <http://www.bbc.co.uk/news/health-27113258> [Accessed 20 July 2014].
- Béland, D. and Cox, R.H. eds., 2011. *Ideas and Politics in Social Science Research*. Oxford: Oxford University Press.
- Béland, D. and Cox, R.H., 2011a. Introduction: Ideas and Politics. In: D. Béland and R.H. Cox, eds., 2011. *Ideas and Politics in Social Science Research*. Oxford: Oxford University Press. Ch. 1
- Béland, D. and Cox, R.H., 2013. The Politics of Policy Paradigms. *Governance: An International Journal of Policy, Administration, and Institutions*, 26(2), pp. 193-195.
- Bergeron, H. and Kopp, P., 2002. Policy Paradigms, Ideas and Interests: The case of the French Public Health Policy toward Drug Abuse. *Annals of the American Academy of Political and Social Science*, 582, pp. 37-48.
- Biller-Andorno, N., Lie, R. and Meulen, R., 2002. Evidence-Based Medicine as an Instrument for Rational Health Policy. *Health Care Analysis*, 10, pp. 261-275.
- Blank, R. and Burau V., 2006. Setting Health Priorities across Nations: More Convergence than Divergence? *Journal of Public Health Policy*, 27(3), pp. 265-281.
- Blank, R. H. and Burau, V., 2004. *Comparative Health Policy*. New York: Palgrave Macmillan.

Boehmer-Christiansen, S. and Skea, J. 1991. *Acid Politics: Environmental and Energy Policies in Britain and Germany*. London: Bellhaven.

Boehmer-Christiansen, S. and Weidner, H. 1995. *The Politics of Reducing Vehicle Emissions in Britain and Germany*. London: Printer in association with the Anglo-German Foundation for the Study of Industrial Society.

Böhm, K. and Landwehr, C., 2013. The Europeanization of Health Care Coverage Decisions: EU-Regulation, Policy Learning and Cooperation in Decision-Making. *Journal of European Integration*, pp. 1-19.

Boseley, S., 2012. Cancer drug 'too expensive for NHS'. *The Guardian*, [online] (Last updated 2<sup>nd</sup> February 2012) Available at: <http://www.theguardian.com/society/2012/feb/02/cancer-drug-too-expensive-nhs> [Accessed 22 July 2014].

Bundesgesetzblatt, 2010. *Gesetz zur Neuordnung des Arzneimittelmarktes in der gesetzlichen Krankenversicherung (Arzneimittelmarktneuordnungsgesetz – AMNOG)*. [online] Available at: [http://www.bgbl.de/Xaver/start.xav?startbk=Bundesanzeiger\\_BGBI&jumpTo=bgbl110s2262.pdf](http://www.bgbl.de/Xaver/start.xav?startbk=Bundesanzeiger_BGBI&jumpTo=bgbl110s2262.pdf) [Accessed 22 November 2013].

Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM), 2013. *Arzneimittelzulassung*. [online] Available at: [http://www.bfarm.de/DE/Arzneimittel/zul/\\_node.html](http://www.bfarm.de/DE/Arzneimittel/zul/_node.html) [Accessed 10 November 2013].

Bundesministerium der Justiz und für Verbraucherschutz, 1949. *Grundgesetz für die Bundesrepublik Deutschland*. [online] Available at: <http://www.gesetze-im-internet.de/bundesrecht/gg/gesamt.pdf> [Accessed 19 December 2013].

Bundesministerium der Justiz und für Verbraucherschutz, 2011. *Verordnung über die Nutzenbewertung von Arzneimittel nach § 35a Absatz 1 SGB V für die Erstattungsvereinbarungen nach § 130b SGB V (AM-NutzenV)*. [online] Available at: <http://www.gesetze-im-internet.de/bundesrecht/am-nutzenv/gesamt.pdf> [Accessed 10 Jul 2013].

Bundesministerium der Justiz und für Verbraucherschutz, 2013. *Sozialgesetzbuch (SGB Fünftes Buch (V)) – Gesetzliche Krankenversicherung*. [online] Available at: [http://www.gesetze-im-internet.de/sgb\\_5/\\_3.html](http://www.gesetze-im-internet.de/sgb_5/_3.html) [Accessed 22 November 2013].

Bundesministerium für Gesundheit (BMG), 2013c. *Gesundheitsfonds*. [online] Available at: <http://www.bmg.bund.de/krankenversicherung/finanzierung/gesundheitsfonds.html> [Accessed 22 November 2013].

Bundesministerium für Gesundheit, 2013a. *Beitragssatz der gesetzlichen Krankenversicherung*. [online] Available at: <http://www.bmg.bund.de/krankenversicherung/beitraege-sozialausgleich/beitragssatz.html> [Accessed 22 November 2013].

Bundesministerium für Gesundheit (BMG), 2013. *Versicherte*. [online] Available at: <http://www.bmg.bund.de/krankenversicherung/versicherte.html> [Accessed 22 November 2013].

Bundesministerium für Gesundheit (BMG), 2013b. *Aufgaben und Organisation der GKV*. [online] Available at: <http://www.bmg.bund.de/krankenversicherung/grundprinzipien/aufgaben-und-organisation-der-gkv.html> [Accessed 22 November 2013].

Busse, R., 1999. Priority-setting and rationing in German health care. *Health Policy*, 50, pp. 71-90.

Busse, R., Schreyoegg, J. and Henke, K-D., 2005. Regulation of pharmaceutical markets in Germany: improving efficiency and controlling expenditures? *International Journal of Health Planning and Management*, 20, pp. 329-349.

Buyx, A.M, 2008. Personal responsibility for health as a rationing criterion: why we don't like it and why maybe we should. *Journal of Medical Ethics*, 34(12), pp. 871-874.

Cabinet Office, 2010. *The Coalition: our programme for government*. [online] Available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/78977/coalition\\_programme\\_for\\_government.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/78977/coalition_programme_for_government.pdf) [Accessed 5 September 2014].

Campbell, J.L., 2002. Ideas, Politics and Public Policy. *Annual Review of Sociology*, 28, pp. 21-38.

Caro, J. J., Nord, E., Siebert, U., McGuire, A., McGregor, M., Henry, D., de Puvourville, G., Attela, V. and Kolominsky-Rabas, P., 2010. The Efficiency Frontier Approach to Economic Evaluation of Health-Care Interventions. *Health Economics*, 19(10), pp. 1117-1127.

Cartwright, N. and Hardie, J., 2012. *Evidence-Based Policy. A Practical Guide to Doing it Better*. Oxford: Oxford University Press.

Clark, S. and Weale, A., 2012. Social values in health priority setting: a conceptual framework. *Journal of Health Organization and Management*, 26(3), pp. 293-316.

Cook, C., 2014. NHS England push to rein in cancer drug prices. BBC, [online] (Last updated 22:01 on 27<sup>th</sup> August 2014) Available at: <http://www.bbc.co.uk/news/health-28938022> (Accessed 28 August 2014).

Coulter, A. and Ham, C. eds, 2000. *The Global Challenge of Health Care Rationing*. Buckingham: Open University Press.

Crosson, S., 2013. The impact of empowering scientific advisory committees to constrain catch limits in US fisheries. *Science and Public Policy*, 40, pp. 261-273.

Daniels, N., 2000. Accountability for reasonableness: Establishing a fair process for priority setting is easier than agreeing on principles. *British Medical Journal*, 321, pp. 1300-1301.

Das Bundesverfassungsgericht, 2005. *Leitsatz zum Beschluss des Ersten Senats vom 6. Dezember 2005 – 1 BvR 347/98* -. [online] Available at: [http://www.bverfg.de/entscheidungen/rs20051206\\_1bvr034798.html](http://www.bverfg.de/entscheidungen/rs20051206_1bvr034798.html) [Accessed 17 December 2013].

Della Porta, D. and Keating, M. eds., 2008. *Approaches and Methodologies in the Social Sciences: A Pluralist Perspective*. Cambridge: Cambridge University Press.

Della Porta, D., 2008. Comparative Analysis: Case-oriented versus Variable-oriented Research. In: D. della Porrrta and M. Keating, eds., 2008. *Approaches and Methodologies in the Social Sciences: A Pluralist Perspective*. Cambridge: Cambridge University Press.

Department of Health. 2009. *Introduction to pharmaceutical price regulation*. [online] Available at: [http://www.dh.gov.uk/en/Healthcare/Medicinespharmacyandindustry/Pharmaceuticalpri ceregulationscheme/DH\\_4071841](http://www.dh.gov.uk/en/Healthcare/Medicinespharmacyandindustry/Pharmaceuticalpri ceregulationscheme/DH_4071841) [Accessed 8 December 2010].

Department of Health, 2013. *Guide to the Healthcare System in England*. [online] Available at: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/194002/9421-2900878-TSO-NHS\\_Guide\\_to\\_Healthcare\\_WEB.PDF](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/194002/9421-2900878-TSO-NHS_Guide_to_Healthcare_WEB.PDF) [Accessed 14 November 2013].

Department of Health, 2013a. *The NHS Constitution*. [online] Available at: <http://www.nhs.uk/choiceintheNHS/Rightsandpledges/NHSConstitution/Documents/2013/the-nhs-constitution-for-england-2013.pdf> [Accessed 3 January 2014].

Deutsche Apotheker Zeitung, 2014. *Kassen fürchten Milliardenkosten*. [online] Available at: <http://www.deutsche-apotheker-zeitung.de/wirtschaft/news/2014/08/06/kassen-fuerchten-milliardenkosten/13507.html> [Accessed 15 September 2014].

Deutsche Apotheker Zeitung, 2014a. *Spahn warnt Gilead: Keine Mondpreise*. [online] Available at: <http://www.deutsche-apotheker-zeitung.de/politik/news/2014/08/07/spahn-warnt-gilead-keine-mondpreise/13521.html> [Accessed 15 September 2014].

Dixon, A. and Poteliakhoff, E., 2012. Back to the future: 10 years of European health reforms. *Health Economics, Policy and Law*, 7(1), pp. 1-10.

Drummond, M. and Rutten, F., 2008. *New Guidelines for Economic Evaluation in Germany and the United Kingdom: Are we any closer to International Standards?* London: Office of Health Economics.

Drummond, M.F., Sculpher, M.J., Torrance, G.W., O'Brien, B.J. and Stoddart, G.L. eds., 2005. *Methods for the Economic Evaluation of Health Care Programmes*. 3rd ed. Oxford: Oxford University Press.

Dunleavy, P., 1981. Professions and policy change: notes towards a model of ideological corporatism. *Public Administration Bulletin*, 36, pp. 3-16.

Edgar, J., 2012. Nice approves prostate cancer drug Abirateron. *The Independent*, [online] (Last updated 16 May 2012) Available at: <http://www.independent.co.uk/life-style/health-and-families/health-news/nice-approves-prostate-cancer-drug-abirateron-7756667.html?origin=internalSearch> [Accessed 28 July 2014].

Eichler, H. G., Bloechi-Daum, B., Abadie, E., Barnett, D., Koenig, F. and Pearson, S., 2010. OUTLOOK Relative Efficacy of Drugs: An Emerging Issue between Regulatory Agencies and Third-party Payers. *Nature Reviews Drug Discovery*, 9 (4), pp. 277-291.

Epilepsy Action, 2012. *What is epilepsy?* | *Epilepsy Action*. [online] Available at: <http://www.epilepsy.org.uk/info/what-is-epilepsy> [Accessed 4 July 2013].

EUnetHTA, 2014. *What is Health Technology Assessment (HTA)*. [online] Available at: <http://www.eunetha.eu/about-us/faq#t287n73> [Accessed 20 May 2014].

European Medicines Agency (EMA), 2013. *What we do*. [online] Available at: [http://www.ema.europa.eu/ema/index.jsp?curl=pages/about\\_us/general/general\\_content\\_000091.jsp&mid=WC0b01ac0580028a42](http://www.ema.europa.eu/ema/index.jsp?curl=pages/about_us/general/general_content_000091.jsp&mid=WC0b01ac0580028a42) [Accessed 10 November 2013].

European Medicines Agency, 2011. *Gilenya: EPAR - Summary for the public*. [online] Available at: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/EPAR\\_-\\_Summary\\_for\\_the\\_public/human/002202/WC500104530.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/002202/WC500104530.pdf) [Accessed 18 June 2013].

European Medicines Agency, 2011a. *European Medicines Agency - Find medicine - Incivo*. [online] Available at: [http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/002313/human\\_med\\_001487.jsp&mid=WC0b01ac058001d124](http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/002313/human_med_001487.jsp&mid=WC0b01ac058001d124) [Accessed 14 June 2013].

European Medicines Agency, 2013. *European Medicines Agency recommends restricting Trobalt to last-line therapy in partial epilepsy*. [online] Available at: [http://www.ema.europa.eu/ema/index.jsp?curl=pages/news\\_and\\_events/news/2013/05/news\\_detail\\_001802.jsp&mid=WC0b01ac058001d126](http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2013/05/news_detail_001802.jsp&mid=WC0b01ac058001d126) [Accessed 7 September 2014].

Federal Joint Committee (FJC), 2013. *Members*. [online] Available at: <http://www.english.g-ba.de/structure/members/> [Accessed 14 November 2013].

Federal Joint Committee (FJC), 2013a. *The Benefit Assessment of Pharmaceuticals in Accordance with the German Social Code Book, Book V (SGB V), section 35a*. [online] Available at: <http://www.english.g-ba.de/benefitassessment/information/> [Accessed 14 November 2013].

Frankfurter Allgemeine Zeitung (F.A.Z.), 2013. *Krankenkassen machen riesigen Überschuss*. [online] Available at: <http://www.faz.net/aktuell/wirtschaft/wirtschaftspolitik/gesetzliche-versicherung-krankenkassen-machen-riesigen-ueberschuss-12551732.html> [Accessed 24 November 2013].

Freeman, R. and Moran, M., 2000. Reforming health care in Europe. *West European Politics*, 23(2), pp. 35-58.

Freeman, R., 2009. Pharmaceutical Policy and Politics in OECD Countries. In: T. R. Marmor, R. Freeman and K. G. H. Okma, eds. 2009. *Comparative Studies and the Politics of Modern Medical Care*. New Haven: Yale University Press. Ch. 9

Gemeinsamer Bundesausschuss, 2012. *Zusammenfassende Dokumentation über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Boceprevir*. [online] Available at: [http://www.g-ba.de/downloads/40-268-1902/2012-03-01\\_AM-RL-XII\\_Boceprevir\\_ZD.pdf](http://www.g-ba.de/downloads/40-268-1902/2012-03-01_AM-RL-XII_Boceprevir_ZD.pdf) [Accessed 6 January 2014].

Gemeinsamer Bundesausschuss, 2012a. *Zusammenfassende Dokumentation über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Telaprevir*. [online] Available at: [http://www.g-ba.de/downloads/40-268-1949/2012-03-29\\_AM-RL-XII\\_Telaprevir\\_ZD.pdf](http://www.g-ba.de/downloads/40-268-1949/2012-03-29_AM-RL-XII_Telaprevir_ZD.pdf) [Accessed 14 June 2013].

Gemeinsamer Bundesausschuss, 2012b. *Zusammenfassende Dokumentation über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Apixaban*. [online] Available at: [http://www.g-ba.de/downloads/40-268-2033/2012-06-07\\_AM-RL-XII\\_Apixaban\\_ZD.pdf](http://www.g-ba.de/downloads/40-268-2033/2012-06-07_AM-RL-XII_Apixaban_ZD.pdf) [Accessed 5 January 2014].

Gemeinsamer Bundesausschuss, 2012c. *Zusammenfassende Dokumentation über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Cabazitaxel*. [online] Available at: [https://www.g-ba.de/downloads/40-268-1974/2012-03-29\\_AM-RL-XII\\_Cabazitaxel\\_ZD.pdf](https://www.g-ba.de/downloads/40-268-1974/2012-03-29_AM-RL-XII_Cabazitaxel_ZD.pdf) [Accessed 14 June 2013].

Gemeinsamer Bundesausschuss, 2012d. *Zusammenfassende Dokumentation über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Eribulin*. [online] Available at: [https://www.g-ba.de/downloads/40-268-2008/2012-04-19\\_AM-RL-XII\\_Eribulin\\_ZD.pdf](https://www.g-ba.de/downloads/40-268-2008/2012-04-19_AM-RL-XII_Eribulin_ZD.pdf) [Accessed 5 March 2014].

Gemeinsamer Bundesausschuss, 2012e. *Zusammenfassende Dokumentation über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Ipilimumab*. [online] Available at: [https://www.g-ba.de/downloads/40-268-2108/2012-08-02\\_AM-RL-XII\\_Ipilimumab\\_ZD.pdf](https://www.g-ba.de/downloads/40-268-2108/2012-08-02_AM-RL-XII_Ipilimumab_ZD.pdf) [Accessed 25 August 2014].

Gemeinsamer Bundesausschuss, 2012f. *Zusammenfassende Dokumentation über die*

*Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Retigabin.* [online] Available at: [https://www.g-ba.de/downloads/40-268-1975/2012-05-03\\_AM-RL-XII\\_Retigabin\\_ZD.pdf](https://www.g-ba.de/downloads/40-268-1975/2012-05-03_AM-RL-XII_Retigabin_ZD.pdf) [Accessed 18 June 2013].

Gemeinsamer Bundesausschuss, 2012g. *Zusammenfassende Dokumentation über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Fingolimod.* [online] Available at: [http://www.g-ba.de/downloads/40-268-1957/2012-06-21\\_AM-RL-XII\\_Fingolimod\\_ZD.pdf](http://www.g-ba.de/downloads/40-268-1957/2012-06-21_AM-RL-XII_Fingolimod_ZD.pdf) [Accessed 18 June 2013].

Gemeinsamer Bundesausschuss, 2012h. *Zusammenfassende Dokumentation über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Abirateronacetat.* [online] Available at: [https://www.g-ba.de/downloads/40-268-1951/2012-03-29\\_AM-RL-XII\\_Abirateron\\_ZD.pdf](https://www.g-ba.de/downloads/40-268-1951/2012-03-29_AM-RL-XII_Abirateron_ZD.pdf) [Accessed 10 September 2014].

Gemeinsamer Bundesausschuss, 2013. *Verfahrensordnung des Gemeinsamen Bundesausschusses.* [online] Available at: [http://www.g-ba.de/downloads/62-492-711/VerfO\\_2013-06-11.pdf](http://www.g-ba.de/downloads/62-492-711/VerfO_2013-06-11.pdf) [Accessed 3 July 2013].

Gemeinsamer Bundesausschuss, 2013a. *Arbeitsweise.* [online] Available at: <http://www.g-ba.de/institution/aufgabe/arbeitsweise/> [Accessed 19 December 2013].

Gemeinsamer Bundesausschuss, 2014. *Beschluss des Gemeinsamen Bundesausschusses über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Sofosbuvir.* [online] Available at: [https://www.g-ba.de/downloads/39-261-2029/2014-07-17\\_AM-RL-XII\\_Sofosbuvir\\_2014-02-01-D-091\\_BAnz.pdf](https://www.g-ba.de/downloads/39-261-2029/2014-07-17_AM-RL-XII_Sofosbuvir_2014-02-01-D-091_BAnz.pdf) [Accessed 25 September 2014].

George, A.L. and Bennett, A., 2005. *Case Studies and Theory Development in the Social Sciences.* Cambridge: The MIT Press.

Giaimo, S. and Manow, P. 1999. Adapting the Welfare State: The Case of Health Care Reform in Britain, Germany and the United States. *Comparative Political Studies*, 32(8), pp. 976-1000.

Giaimo, S., 1995. Health Care Reform in Britain and Germany: Recasting the Political Bargain with the Medical Profession. *Governance*, 8(3), pp. 354-379.

Giaimo, S., 2001. Who pays for Health Care Reform? In: P. Pierson, ed., 2001. *The New Politics of the Welfare State.* Oxford: Oxford University Press.

Goldenberg, M.J., 2006. On evidence and evidence-based medicine: Lessons from the philosophy of science. *Social Science & Medicine*, 62, pp. 2621-2632.

Guston, D.H., 1996. Principal-agent theory and the structure of science policy. *Science and Public Policy*, 23(4), pp. 229-240.

- Haas, P., 1992. Introduction: epistemic communities and international policy coordination. *International Organization*, 46(1), pp. 1-33.
- Hall, P., 1993. Policy Paradigms, Social Learning, and the State: the Case of Economic Policymaking in Britain. *Comparative Politics*, 25(3), p. 275-296.
- Hall, P.A., 1989. *The Political Power of Economic Ideas: Keynesianism across Nations*. Princeton: Princeton University Press.
- Hall, P.A., 2013. Brother, Can You Paradigm? *Governance: An International Journal of Policy, Administration, and Institutions*, 26(2), pp. 189-192.
- Ham, C. and Honigsbaum, F., 1998. Priority setting and rationing health services. In: R.B. Saltman, J. Figueras and C. Sakellarides, eds. 1998. *Critical Challenges for Health Care Reform in Europe*. Maidenhead: Open University Press. Ch. 4
- Ham, C. and Robert, G. eds., 2003. *Reasonable Rationing: International Experience of Priority Setting in Health Care*. Maidenhead: Open University Press.
- Ham, C. ed., 1997. *Health Care Reform: Learning from International Experience* Buckingham: Open University Press.
- Ham, C., 2009. *Health Policy in Britain*. Basingstoke: Palgrave Macmillan.
- Hanisch, M. and Kanavos, P., 2008. Editorial – cost containment: impact and consequences. *Eurohealth*, [online] Available at: [http://www.lse.ac.uk/LSEHealthAndSocialCare/pdf/eurohealth/VOL14N2/Eurohealth%20summer%202008\\_temp%20web.pdf](http://www.lse.ac.uk/LSEHealthAndSocialCare/pdf/eurohealth/VOL14N2/Eurohealth%20summer%202008_temp%20web.pdf) [Accessed 23 August 2014].
- Harrison, S., Moran, M. and Wood, B., 2002. Policy emergence and policy convergence: the case of ‘scientific-bureaucratic medicine’ in the United States and United Kingdom. *The British Journal of Politics and International Relations*, 4(1), pp. 1-24.
- Hawe, E. and Cockcroft, L., 2013. *OHE Guide to UK Health and Health Care Statistics Second Edition*. [pdf] London: Office of Health Economics. Available at: <https://www.ohe.org/publications/ohe-guide-uk-health-and-health-care-statistics> [Accessed 25 October 2014].
- Holm, S., 1998. Goodbye to the simple solutions: the second phase of priority setting in health care. *British Medical Journal*, 317, pp. 1000-1002.
- Honig, M.I., 2006. Street-level Bureaucracy Revisited: Frontline District Central-Office Administrators as Boundary Spanners in Education Policy Implementation. *Educational Evaluation and Policy Analysis*, 28(4), pp. 357-383.
- Hope, J., 2012. Now patients are denied ‘breakthrough’ prostate cancer drug as health watchdogs brand it too expensive. *The Daily Mail*, [online] (Last updated 00.04 on 2<sup>nd</sup> February 2012) Available at: <http://www.dailymail.co.uk/health/article-2095087/Prostate-cancer-patients-denied-expensive-drug.html> [Accessed 22 July 2014].

HTA Glossary, 2014. *Health technology assessment (HTA)*. [online] Available at: <http://htaglossary.net/health+technology+assessment+%28HTA%29> [Accessed 24 August 2014].

HTA Glossary, 2014a. *Subgroup analysis*. [online] Available at: <http://htaglossary.net/subgroup+analysis> [Accessed 7 September 2014].

Immergut, E., 1992. *Health Politics: Interests and Institutions in Western Europe*. Cambridge: Cambridge University Press.

Institut der deutschen Wirtschaft Köln and vfa Die forschenden Pharma-Unternehmen, 2013. *Pharmaindustrie in Deutschland*. [pdf] Köln: Institut der deutschen Wirtschaft Köln and vfa Die forschenden Pharma-Unternehmen. Available at: [www.pharmastandort.de/.../Pharmaindustrie\\_Deutschland\\_2012.pdf](http://www.pharmastandort.de/.../Pharmaindustrie_Deutschland_2012.pdf) [Accessed 25 October 2014].

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2012e. *Retigabine – Benefit assessment according to § 35a Social Code Book V*. [online] Available at: [https://www.iqwig.de/download/A11-28\\_Retigabine\\_Extract\\_of\\_dossier\\_assessment.pdf](https://www.iqwig.de/download/A11-28_Retigabine_Extract_of_dossier_assessment.pdf) [Accessed 18 December 2013].

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2011a. *Allgemeine Methoden*. [online] Available at: [https://www.iqwig.de/download/IQWiG\\_Methoden\\_Version\\_4\\_0.pdf](https://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf) [Accessed 8 November 2013].

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2012a. *Cabazitaxel - Nutzenbewertung gemäß § 35a SGB V*. [online] Available at: [https://www.iqwig.de/download/A11-24\\_Cabazitaxel\\_Nutzenbewertung\\_35a\\_SGB\\_V.pdf](https://www.iqwig.de/download/A11-24_Cabazitaxel_Nutzenbewertung_35a_SGB_V.pdf) [Accessed 24 May 2013].

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2012b. *Eribulin - Nutzenbewertung gemäß § 35a SGB V*. [online] Available at: [https://www.iqwig.de/download/A11-26\\_Eribulin\\_Nutzenbewertung\\_35a\\_SGB\\_V.PDF](https://www.iqwig.de/download/A11-26_Eribulin_Nutzenbewertung_35a_SGB_V.PDF) [Accessed 4 September 2014].

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2012c. *Ipilimumab - Nutzenbewertung gemäß § 35a SGB V*. [online] Available at: [https://www.iqwig.de/download/A12-07\\_Ipilimumab\\_Nutzenbewertung\\_35a\\_SGB\\_V.pdf](https://www.iqwig.de/download/A12-07_Ipilimumab_Nutzenbewertung_35a_SGB_V.pdf) [Accessed 23 June 2013].

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2012d. *Telaprevir – Nutzenbewertung gemäß § 35a SGB V*. [online] Available at: [https://www.iqwig.de/download/A11-25\\_Telaprevir\\_Nutzenbewertung\\_35a\\_SGB\\_V.PDF](https://www.iqwig.de/download/A11-25_Telaprevir_Nutzenbewertung_35a_SGB_V.PDF) [Accessed 10 June 2013].

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2011b. *Boceprevir – Nutzenbewertung gemäß § 35a SGB V*. [online] Available at: [https://www.iqwig.de/download/A11-17\\_Boceprevir\\_Nutzenbewertung\\_gemaess\\_35a\\_SGB\\_V.pdf](https://www.iqwig.de/download/A11-17_Boceprevir_Nutzenbewertung_gemaess_35a_SGB_V.pdf) [Accessed 10 June 2013].

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen 2011. *Ticagrelor – Nutzenbewertung gemäß § 35a SGB V*. [online] Available at: [https://www.iqwig.de/download/A11-02\\_Ticagrelor\\_Nutzenbewertung\\_§35a\\_SGB\\_V\\_.pdf](https://www.iqwig.de/download/A11-02_Ticagrelor_Nutzenbewertung_§35a_SGB_V_.pdf) [Accessed 07 November 2011].

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2012. *Retigabin - Nutzenbewertung gemäß § 35a SGB V*. [online] Available at: [https://www.g-ba.de/downloads/92-975-62/2011-05-15-D-008\\_Retigabin\\_IQWiG-Nutzenbewertung.pdf](https://www.g-ba.de/downloads/92-975-62/2011-05-15-D-008_Retigabin_IQWiG-Nutzenbewertung.pdf) [Accessed 19 June 2013].

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2012e. *Fingolimod - Nutzenbewertung gemäß § 35a SGB V*. [online] Available at: [https://www.iqwig.de/download/A11-23\\_Fingolimod\\_Nutzenbewertung\\_35a\\_SGB\\_V.pdf](https://www.iqwig.de/download/A11-23_Fingolimod_Nutzenbewertung_35a_SGB_V.pdf) [Accessed 19 June 2013].

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, 2011c. *Abirateronacetat - Nutzenbewertung gemäß § 35a SGB V*. [online] Available at: [https://www.g-ba.de/downloads/92-975-26/2011-10-01-D-023\\_Abirateronacetat\\_IQWiG-Nutzenbewertung.PDF](https://www.g-ba.de/downloads/92-975-26/2011-10-01-D-023_Abirateronacetat_IQWiG-Nutzenbewertung.PDF) [Accessed 10 September 2014].

Institute for Quality and Efficiency in Health Care (IQWiG), 2013. *Responsibilities and Objectives of IQWiG*. [online] Available at: [https://www.iqwig.de/en/about\\_us/responsibilities\\_and\\_objectives\\_of\\_iqwig.2946.html](https://www.iqwig.de/en/about_us/responsibilities_and_objectives_of_iqwig.2946.html) [Accessed 15 November 2013].

Institute for Quality and Efficiency in Health Care, 2008. *Methods for the Assessment of the Relation of Benefits to Costs in the German Statutory Health Care System*. [online] Available at: [https://www.iqwig.de/download/08-01-24\\_Draft\\_Methods\\_of\\_the\\_Relation\\_of\\_Benefits\\_to\\_Costs\\_Version\\_1\\_0.pdf](https://www.iqwig.de/download/08-01-24_Draft_Methods_of_the_Relation_of_Benefits_to_Costs_Version_1_0.pdf) [Accessed 13 November 2011].

Interviewee No. 1, 2012. *The technology appraisal of Telaprevir and Boceprevir*. Interviewed by Katharina Kieslich [in person] London, 6 October 2012.

Interviewee No. 10, 2013. *The technology appraisal of Apixaban*. Interviewed by Katharina Kieslich [in person] London, 26 March 2013.

Interviewee No. 11, 2013. *The technology appraisal of pharmaceuticals in England*. Interviewed by Katharina Kieslich [in person] London, 3 April 2013.

Interviewee No. 12, 2013. *The early benefit assessment of Ticagrelor*. Interviewed by Katharina Kieslich [in person] Hamburg, 9 April 2013.

Interviewee No. 13, 2013. *The early benefit assessment of Retigabine*. Interviewed by Katharina Kieslich [in person] Munich, 11 April 2013.

Interviewee No. 14, 2013. *The early benefit assessment of Fingolimod*. Interviewed by Katharina Kieslich [in person] Nuremberg, 15 April 2013.

Interviewee No. 15. 2013. *The early benefit assessment of Retigabine in Germany*. Interviewed by Katharina Kieslich [in person] Erlangen, 15 April 2013.

Interviewee No. 16, 2013. *The early benefit assessment process in Germany*. Interviewed by Katharina Kieslich [in person] Berlin, 26 April 2013.

Interviewee No. 17, 2013. *The early benefit assessment of Fingolimod*. Interviewed by Katharina Kieslich [in person] Berlin, 26 April 2013.

Interviewee No. 18, 2013. *The technology appraisal of Eribulin*. Interviewed by Katharina Kieslich [in person] Sheffield, 7 May 2013.

Interviewee No. 19, 2013. *The technology appraisal of Abiraterone and Cabazitaxel*. Interviewed by Katharina Kieslich [in person] London, 8 May 2013.

Interviewee No. 2, 2012. *The technology appraisal of Telaprevir*. Interviewed by Katharina Kieslich [in person] London, 31 October 2012.

Interviewee No. 20, 2013. *The technology appraisal of Retigabine*. Interviewed by Katharina Kieslich [in person] London, 9 May 2013.

Interviewee No. 21, 2013. *The technology appraisal of Apixaban*. Interviewed by Katharina Kieslich [in person] London, 29 May 2013.

Interviewee No. 22, 2013. *The technology appraisal of Fingolimod*. Interviewed by Katharina Kieslich [in person] London, 17 June 2013.

Interviewee No. 23, 2013. *The early benefit assessment process in Germany*. Interviewed by Katharina Kieslich [in person] Cologne, 24 June 2013.

Interviewee No. 3, 2012. *The technology appraisal of Telaprevir*. Interviewed by Katharina Kieslich [in person] London, 2 November 2012.

Interviewee No. 4, 2013. *The technology appraisal of Eribulin*. Interviewed by Katharina Kieslich [in person] London, 7 March 2013.

Interviewee No. 5, 2013. *The early benefit assessment of Telaprevir and Abiraterone*. Interviewed by Katharina Kieslich [in person] Neuss, 14 February 2013.

Interviewee No. 6, 2013. *The early benefit assessment process in Germany*. Interviewed by Katharina Kieslich [in person] Berlin, 6 March 2013.

Interviewee No. 7, 2013. *The early benefit assessment of Abiraterone, Cabazitaxel, Eribulin and Ipilimumab*. Interviewed by Katharina Kieslich [in person] Berlin, 6 March 2013.

Interviewee No. 8, 2013. *The early benefit assessment process in Germany*. Interviewed by Katharina Kieslich [in person] Berlin, 7 March 2013.

Interviewee No. 9, 2013. *The early benefit assessment of Cabazitaxel*. Interviewed by Katharina Kieslich [in person] Berlin, 7 March 2013.

Kanavos, P., Nicod, E., van den Aardweg, S. and Pomedli, S., 2010. The Impact of

Health Technology Assessments: An International Comparison. *Euro Observer: The Health Policy Bulletin of the European Observatory on Health Systems and Policies*, 12(4), pp. 1-7.

Kay, A., 2007. Tense Layering and Synthetic Policy Paradigms: The Politics of Health Insurance in Australia. *Australian Journal of Political Science*, 42(4), pp. 579-591.

Kenny, N. and Joffres, C., 2008. An Ethical Analysis of International Health Priority-setting. *Health Care Analysis*, 16(2), pp. 145-160.

Klein, R., 2003. Evidence and policy: interpreting the Delphic oracle. *Journal of the Royal Society of Medicine*, 96, pp. 429-431.

Klein, R., 2010. Rationing in the fiscal ice age. *Health Economics, Policy and Law*, 5, pp. 389-396.

Klingler, C., Shah, S.M.B, Barron, A.J.G and Wright, J.S.F., 2013. Regulatory space and the contextual mediation of common functional pressures: Analyzing the factors that led to the German Efficiency Frontier approach. *Health Policy*, 109, pp. 270-280.

Kuhn, T.S. 1969. Postscript. In: T.S. Kuhn, 1962. *The Structure of Scientific Revolutions*. 2<sup>nd</sup> ed. Chicago: The University of Chicago Press.

Kuhn, T.S., 1962. *The Structure of Scientific Revolutions*. 2<sup>nd</sup> ed. Chicago: The University of Chicago Press.

Kurzer, P., 2013. The Politics of Ideas in Reforming the Dutch Disability Fund. *Governance: An International Journal of Policy, Administration, and Institutions*, 26(2), pp. 283-305.

Landwehr, C., 2009. Deciding how to decide: the case of health care rationing. *Public Administration*, 87(3), pp. 586-603.

Lehoux, P. and Blume, S., 2000. Technology Assessment and the Sociopolitics of Health Technologies. *Journal of Health Politics, Policy and Law*, 25(6), pp. 1083-1120.

Littlejohns, P., Weale, A., Chalkidou, K., Faden, R. and Teerawattananon, Y., 2012. Social values and health policy: a new international research programme. *Journal of Health Organization and Management*, 26(3), pp. 285-292.

Lowi, T.J., 1964. American Business, Public Policy, Case-Studies and Political Theory. *World Politics*, 16(4), pp. 677-715.

Majone, G., 1989. *Evidence, Argument, & Persuasion in the Policy Process*. New Haven and London: Yale University Press.

Marckmann, G., 2010. Kann Rationierung im Gesundheitswesen ethisch vertretbar sein? *G&G Wissenschaft*, [online] Available at: [http://www.wido.de/fileadmin/wido/downloads/pdf\\_ggw/wido\\_ggw\\_aufs1\\_0110.pdf](http://www.wido.de/fileadmin/wido/downloads/pdf_ggw/wido_ggw_aufs1_0110.pdf) [Accessed 27 August 2014].

Marmor, T.R., Freeman, R. and Okma, K.G.H. eds., 2009. *Comparative Studies and the Politics of Modern Medical Care*. New Haven: Yale University Press.

May, P.J. and Winter, S.C., 2009. Politicians, Managers, and Street-Level Bureaucrats: Influences on Policy Implementation. *Journal of Public Administration Research and Theory*, 19(3), pp. 453-476.

Maynard, A. and Bloor, K., 2003. Dilemmas in regulation of the market for pharmaceuticals. *Health Affairs*, 22(3), pp. 31-41.

McKinnell, C., 2011. *Multiple sclerosis (relapsing-remitting) – Fingolimod*. [online] Available at: <https://www.nice.org.uk/guidance/ta254/resources/multiple-sclerosis-relapsingremitting-fingolimod-catherine-mckinnell-mp2> [Accessed 12 September 2014].

Medicines and Healthcare Products Regulatory Agency (MHRA), 2012. *Medicines and Medical Devices Regulation: What you need to know*. [online] Available at: <http://www.mhra.gov.uk/home/groups/comms-ic/documents/websiteresources/con2031677.pdf> [Accessed 10 November 2013].

Melo, M.A., 2004. Institutional Choice and the Diffusion of Policy Paradigms: Brazil and the Second Wave of Pension Reform. *International Political Science Review*, 25(3), pp. 320-341.

Menaheim, G., 1998. Policy Paradigms, Policy Networks and Water Policy in Israel. *Journal of Public Policy*, 18, pp. 283-310.

Menaheim, G., 2008. The Transformation of Higher Education in Israel since the 1990s: The Role of Ideas and Policy Paradigms. *Governance: An International Journal of Policy, Administration, and Institutions*, 21(4), pp. 499-526.

Milewa, T. and Barry, C., 2005. Health Policy and the Politics of Evidence. *Social Policy & Administration*, 39(5), pp. 498-512.

Moran, M., 1995. Three Faces of the Health Care State. *Journal of Health Politics, Policy and Law*, 20(3), pp. 767-781.

Moran, M., 1999. *Governing the health care state. A comparative study of the United Kingdom, the United States and Germany*. Manchester: Manchester University Press.

Mossialos, E. and Oliver, A., 2005. An overview of pharmaceutical policy in four countries: France, Germany, the Netherlands and the United Kingdom. *International Journal of Health Planning and Management*, 20, pp. 291-306.

Mossialos, E., Mrazek, M. and Walley T. eds., 2004. *Regulating pharmaceuticals in Europe: striving for efficiency, equity and equality*. Maidenhead: Open University Press.

Mossialos, E., Walley, T. and Mrazek, M., 2004. Regulating pharmaceuticals in Europe: an overview. In: E. Mossialos, M. Mrazek and T. Walley, eds. 2004. *Regulating pharmaceuticals in Europe: striving for efficiency, equity and equality*. Maidenhead: Open University Press. Ch. 1

Mrazek, M. and Mossialos, E., 2004. Regulating pharmaceutical prices in the European Union. In: E. Mossialos, M. Mrazek and T. Walley, eds. 2004. *Regulating pharmaceuticals in Europe: striving for efficiency, equity and equality*. Maidenhead: Open University Press. Ch. 6

National Institute for Health and Care Excellence, 2008. *Guide to methods of technology appraisal*. [online] Available at: <http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf> [Accessed 15 November 2013].

National Institute for Health and Care Excellence, 2008a. *Social Value Judgements: Principles for the development of NICE guidance*. [online] Available at: <http://www.nice.org.uk/media/C18/30/SVJ2PUBLICATION2008.pdf> [Accessed 10 Jul 2013].

National Institute for Health and Care Excellence, 2009. *Technology Appraisal Committee*. [online] Available at: [http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalcommittee/technology\\_appraisal\\_committee.jsp](http://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalcommittee/technology_appraisal_committee.jsp) [Accessed 14 November 2013].

National Institute for Health and Care Excellence, 2009a. *National Institute for Health and Clinical Excellence: Appraising life-extending, end of life treatments*. [online] Available at: <http://www.nice.org.uk/media/E4A/79/SupplementaryAdviceTACEoL.pdf> [Accessed 6 January 2014].

National Institute for Health and Care Excellence, 2011. *Ticagrelor for the treatment of acute coronary syndromes*. [online] Available at: <http://www.nice.org.uk/guidance/ta236/resources/guidance-ticagrelor-for-the-treatment-of-acute-coronary-syndromes-pdf> [Accessed 23 August 2014].

National Institute for Health and Care Excellence, 2011a. *Retigabine for the adjunctive treatment of partial onset seizures in epilepsy*. [online] Available at: <http://www.nice.org.uk/nicemedia/live/13516/55426/55426.pdf> [Accessed 17 June 2013].

National Institute for Health and Care Excellence, 2011b. *Multiple sclerosis (relapsing-remitting) - Fingolimod: appraisal consultation document*. [online] Available at: <http://guidance.nice.org.uk/TA/Wave20/71/ACD1/ACD1> [Accessed 18 June 2013].

National Institute for Health and Care Excellence, 2012. *Developing NICE technology appraisals*. [online] Available at: [http://www.nice.org.uk/aboutnice/howwework/devnicetech/developing\\_nice\\_technology\\_appraisals.jsp](http://www.nice.org.uk/aboutnice/howwework/devnicetech/developing_nice_technology_appraisals.jsp) [Accessed 15 November 2013].

National Institute for Health and Care Excellence, 2012a. *Apixaban for the prevention*

*of venous thromboembolism after total hip or knee replacement in adults.* [online]. Available at: <http://www.nice.org.uk/nicemedia/live/13648/57895/57895.pdf> [Accessed 5 January 2014].

National Institute for Health and Care Excellence, 2012b. *Cabazitaxel for hormone-refractory metastatic prostate cancer previously treated with a docetaxel-containing regime.* [online] Available at: <http://www.nice.org.uk/guidance/ta255/resources/guidance-cabazitaxel-for-hormonerefractory-metastatic-prostate-cancer-previously-treated-with-a-docetaxelcontaining-regimen-pdf> [Accessed 5 September 2014].

National Institute for Health and Care Excellence, 2012c. *Eribulin for the treatment of locally advanced or metastatic breast cancer.* [online] Available at: <https://www.nice.org.uk/guidance/ta250/resources/guidance-eribulin-for-the-treatment-of-locally-advanced-or-metastatic-breast-cancer-pdf> [Accessed 12 August 2014].

National Institute for Health and Care Excellence, 2012d. *Fingolimod for the treatment of highly active relapsing-remitting multiple sclerosis.* [online] Available at: <http://www.nice.org.uk/nicemedia/live/13719/58914/58914.pdf> [Accessed 18 June 2013].

National Institute for Health and Care Excellence, 2012e. *Telaprevir for the treatment of genotype 1 chronic hepatitis C.* [online] Available at: <http://www.nice.org.uk/nicemedia/live/13717/58912/58912.pdf> [Accessed 15 June 2013].

National Institute for Health and Care Excellence, 2012f. *Abiraterone for castration-resistant metastatic prostate cancer previously treated with a docetaxel-containing regimen.* [online] Available at: <https://www.nice.org.uk/guidance/ta259/resources/guidance-abiraterone-for-castrationresistant-metastatic-prostate-cancer-previously-treated-with-a-docetaxelcontaining-regimen-pdf> [Accessed 12 September 2014].

National Institute for Health and Care Excellence, 2012g. *Prostate cancer (metastatic, castration resistant) – abiraterone (following cytotoxic therapy): appraisal consultation.* [online] Available at: <https://www.nice.org.uk/guidance/ta259/resources/prostate-cancer-metastatic-castration-resistant-abiraterone-following-cytotoxic-therapy-appraisal-consultation2> [Accessed 12 September 2014].

National Institute for Health and Care Excellence, 2013. *Guide to the methods of technology appraisal 2013 – Foreword.* [online] Available at: <http://www.nice.org.uk/article/pmg9/chapter/foreword> [Accessed 9 October 2014].

National Institute for Health and Care Excellence, 2013a. *Who we are.* [online] Available at: [http://www.nice.org.uk/aboutnice/whoweare/who\\_we\\_are.jsp](http://www.nice.org.uk/aboutnice/whoweare/who_we_are.jsp) [Accessed 14 November 2013].

National Institute for Health and Care Excellence, 2014. *Breast cancer drug costing tens of thousands of pounds more than other treatments ‘unaffordable’ for NHS.* [online] Available at: <https://www.nice.org.uk/News/Press-and-Media/breast-cancer->

drug-costing-tens-of-thousands-of-pounds-more-than-other-treatments-unaffordable-for-nhs [Accessed 20 July 2014].

National Institute for Health and Care Excellence, 2014a. *NICE technology appraisal guidance*. [online] Available at: <https://www.nice.org.uk/About/What-we-do/Our-Programmes/NICE-guidance/NICE-technology-appraisal-guidance> [Accessed 25 August 2014].

National Institute for Health and Care Excellence, 2014b. *Glossary*. [online] Available at: <https://www.nice.org.uk/Glossary?letter=Q> [Accessed 27 August 2014].

National Institute for Health and Care Excellence, 2014c. *Final appraisal determination: Ipilimumab for previously untreated advanced (unresectable or metastatic) melanoma*. [online] Available at: <https://www.nice.org.uk/guidance/ta319/resources/melanoma-previously-untreated-unresectable-stage-iii-or-iv-ipilimumab-id74-final-appraisal-determination-document2> [Accessed 5 September 2014].

National Institute for Health and Care Excellence, 2014d. *Kadcyla: NICE disappointed by manufacturer's decision*. [online] Available at: <https://www.nice.org.uk/news/press-and-media/kadcyla-nice-disappointed-by-manufacturers-decision> [Accessed 15 September 2014].

NHS England, 2014. *The Cancer Drugs Fund*. [online] Available at: <http://www.england.nhs.uk/ourwork/pe/cdf/> [Accessed 12 September 2014].

Nord, E., 1999. *Cost-Value Analysis in Health Care: Making Sense of QALYs*. Cambridge: Cambridge University Press.

Norheim, O., 2002. The Role of Evidence in Health Policy Making: A Normative Perspective. *Health Care Analysis*, 10, pp. 309-317.

Office for National Statistics, 2012. *Business Enterprise Research and Development, 2011*. [online] Available at: [http://www.ons.gov.uk/ons/dcp171778\\_287868.pdf](http://www.ons.gov.uk/ons/dcp171778_287868.pdf) [Accessed 26 October 2014].

Orenstein, M.A., 2013. Pension Privatization: Evolution of a Paradigm. *Governance: An International Journal of Policy, Administration, and Institutions*, 26(2), pp. 259-281.

Organisation for Economic Co-operation and Development, 2012. *Health at a Glance: Europe 2012*. [online] Available at: <http://www.oecd-ilibrary.org/sites/9789264183896-en/05/05/index.html?itemId=/content/chapter/9789264183896-55-en> [Accessed 25 October 2014].

Organisation for Economic Co-operation and Development, 2013. *OECD Health Statistics 2013 Frequently Requested Data*. [online] Available at: <http://www.oecd.org/health/health-systems/health-statistics.htm> [Accessed 06 July 2014].

Organisation for Economic Co-operation and Development, 2014. *OECD Health*

*Statistics 2014 Frequently Requested Data.* [online] Available at: <http://www.oecd.org/health/health-systems/health-statistics.htm> [Accessed 24 August 2014].

Parsons, W., 2005. *Public Policy: An Introduction to the Theory and Practice of Policy Analysis.* Cheltenham: Edward Elgar Publishing Limited.

Pawson, R., Wong, G., Owen, L., 2011. Known Knowns, Known Unknowns, Unknowns Unknowns: The Predicament of Evidence-Based Policy. *American Journal of Evaluation*, 32(4), pp. 518-546.

Pfaff, M., 2009. Germany: Evidence, Policy, and Politics in Health Care Reform. In: T.R. Marmor, R. Freeman and K.G.H. Okma, eds., 2009. *Comparative Studies and the Politics of Modern Medical Care.* New Haven: Yale University Press. Ch. 4

Pierson, P. ed., 2001. *The New Politics of the Welfare State.* Oxford: Oxford University Press.

Pomedli, S., 2010. HTA and access to cancer medicines. *Euro Observer: The Health Policy Bulletin of the European Observatory on Health Systems and Policies*, 12(4), pp. 7-9.

Poteete, A., 2003. Ideas, Interests, and Institutions: Challenging Property Rights in Botswana. *Governance: An International Journal of Policy, Administration, and Institutions*, 16(4), pp. 527-557.

Radaelli, C.M., 1995. The role of knowledge in the policy process. *Journal of European Public Policy*, 2(2), pp. 159-183.

Ragin, C.C., 1994. *Constructing Social Research.* Thousand Oaks: Pine Forge Press.

Rawlins, M., 2012. *What constitutes credible evidence of effectiveness? 19<sup>th</sup> Annual Lecture 2012.* London: Office of Health Economics.

Robert, G., 2003. The United Kingdom. In: C. Ham and G. Robert, eds., 2003. *Reasonable Rationing: International Experience of Priority Setting in Health Care.* Maidenhead: Open University Press. Ch. 5

Sabatier, P.A. and Jenkins-Smith, H.C. eds., 1993. *Policy Change and Learning: An Advocacy Coalition Approach.* Boulder: Westview Press.

Sabatier, P.A. and Jenkins-Smith, H.C., 1999. The Advocacy Coalition Framework: An Assessment. In: P.A. Sabatier, ed., 1999. *Theories of the Policy Process.* Oxford: Westview Press. Ch. 6

Sabatier, P.A. ed., 1999. *Theories of the Policy Process.* Oxford: Westview Press.

Sabatier, P.A., 1988. An Advocacy Coalition Framework of Policy Change and the Role of Policy-Oriented Learning Therein. *Policy Sciences*, 21(2), pp. 129-168.

Saltman, R., 2012. The role of comparative health studies for policy learning. *Health Economics, Policy and Law*, 7(1), pp. 11-13.

Saltman, R.B., Figueras, J. and Sakellarides, C. eds., 1998. *Critical Challenges for Health Care Reform in Europe*. Maidenhead: Open University Press.

Seiter, A., 2010. *A Practical Approach to Pharmaceutical Policy*. Washington D.C.: The World Bank.

Simon, M., 2010. *Das Gesundheitssystem in Deutschland: Eine Einführung in die Struktur und Arbeitsweise*. Bern: Verlag Hans Huber.

Sloan, F.A. ed., 1995. *Valuing Health Care: Costs, Benefits, and Effectiveness of Pharmaceuticals and other Medical Technologies*. Cambridge: Cambridge University Press.

Sorenson, C. and Chalkidou, K., 2012. Reflections on the evolution of HTA in Europe. *Health Economics, Policy and Law*, 7(1), pp. 25-45.

Sorenson, C., 2009. The Role of HTA in Coverage and Pricing Decisions: A Cross-Country Comparison. *Euro Observer: The Health Policy Bulletin of the European Observatory on Health Systems and Policies*, 11(1), pp. 1-5.

Statista, 2014. *Anteil der Ausgaben für Arzneimittel aus Apotheken an den Gesamtausgaben der gesetzlichen Krankenversicherung (GKV) von 2002 bis 2013*. [online] Available at: <http://de.statista.com/statistik/daten/studie/150562/umfrage/anteil-der-arzneimittelausgaben-an-den-gesamtausgaben-der-gkv/> [Accessed 25 October 2014].

Strech, D., Danis, M., Loeb, M. and Marckmann, G., 2009. Ausmaß und Auswirkungen von Rationierung in deutschen Krankenhäusern. *Deutsche Medizinische Wochenschrift*, 134, pp. 1261-1266.

Sturm, R. and Wilks, S. 1997. *Competition and the Regulation of the Electricity Supply Industry in Britain and Germany*. London: Anglo-German Foundation for the Study of Industrial Society.

Tansey, O., 2007. Process Tracing and Elite Interviewing: A Case for Non-probability Sampling. *PS: Political Science and Politics*, 4, pp. 765-772.

The Association of the British Pharmaceutical Industry (ABPI), 2013. *Global pharmaceutical industry and market*. [online] Available at: <http://www.abpi.org.uk/industry-info/knowledge-hub/global-industry/Pages/industry-market-.aspx#fig1> [Accessed 25 August 2014].

Van Der Wilt, G.J. and Reuzel, R., 1998. Assessment of Health Technologies. Which Issues Should be Addressed? *Evaluation*, 4(3), pp. 351-358.

Velasco-Garrido, M. and Busse, R. 2005. *Health Technology Assessment and Health Policy-making in Europe: Current Status, Challenges and Potential*. Copenhagen:

European Observatory on Health Systems and Policy.

Walker, P. and Batty, D., 2008. Patients protest after kidney cancer drugs rejected. *The Guardian*, [online] (Last updated 15.35 on 27<sup>th</sup> August 2008) Available at: <http://www.theguardian.com/society/2008/aug/27/health.cancer> [Accessed 20 July 2014].

Walley, T., Mrazek, M. and Mossialos, E., 2005. Regulating pharmaceutical markets: Improving efficiency and controlling costs in the UK. *International Journal of Health Planning and Management*, 20, pp. 375-398.

Weale, A., O’Riordan, T. and Kramme, L. 1991. *Controlling Pollution in the Round: Change and Choice in Environmental Regulation in Britain and West Germany*. London: Anglo-German Foundation.

Weiss, C.H., 1979. The many meanings of research utilization. *Public Administration Review*, pp. 426-431.

Weiss, C.H., 1999. The Interface between Evaluation and Public Policy. *Evaluation*, 5(4), pp. 468-486.

Williams, A., 2004. *What could be nicer than NICE? OHE Annual Lecture 2004*. London: Office of Health Economics.

World Health Organization, 2014. *Health System Strengthening Glossary*. [online] Available at: [http://www.who.int/healthsystems/hss\\_glossary/en/index5.html](http://www.who.int/healthsystems/hss_glossary/en/index5.html) [Accessed 6 January 2014].

Yin, R.K., 2009. *Case Study Research: Design and Methods*. Los Angeles: SAGE Publications.

Zentrale Ethikkommission bei der Bundesärztekammer, 2007. *Priorisierung medizinischer Leistungen im System der gesetzlichen Krankenversicherung (GKV)*. [online] Available at: <http://www.zentrale-ethikkommission.de/page.asp?his=0.1.53> [Accessed 21 November 2013].

## Additional Sources<sup>38</sup>

National Institute for Health and Care Excellence, 2011. *Ticagrelor for the treatment of acute coronary syndromes. Reponse to consultee, commentator and public comments on the Appraisal Consultation Documentation (ACD)*. [online] Available at: <http://www.nice.org.uk/guidance/ta236/documents/acute-coronary-syndromes-ticagrelor-response-to-consultee-commentator-and-public-comments-on-the-appraisal-consultation-document-acd2> [Accessed 30 September 2011].

National Institute for Health and Care Excellence, 2012a. *Abiraterone for castration-resistant metastatic prostate cancer previously treated with a docetaxel-containing regimen. Reponse to consultee, commentator and public comments on the Appraisal Consultation Documentation (ACD)*. [online] Available at: <http://www.nice.org.uk/guidance/ta259/documents/response-to-consultee-commentator-and-public-comments-on-the-appraisal-consultation-document2> [Accessed 5 July 2013].

National Institute for Health and Care Excellence, 2012b. *Hepatitis C (genotype 1) – boceprevir: patient group, professional group and NHS organisation submission statements*. [online] Available at: <http://www.nice.org.uk/guidance/ta253/documents> [Accessed 26 April 2013].

National Institute for Health and Care Excellence, 2012c. *Cabazitaxel for hormone-refractory metastatic prostate cancer previously treated with a docetaxel-containing regimen. Reponse to consultee, commentator and public comments on the Appraisal Consultation Documentation (ACD)*. [online] Available at: <http://www.nice.org.uk/guidance/ta255/documents/prostate-cancer-cabazitaxel-response-to-consultee-commentator-and-public-comments-on-the-appraisal-consultation-document-acd2> [Accessed 5 September 2014].

National Institute for Health and Care Excellence, 2012d. *Eribulin for the treatment of locally advanced or metastatic breast cancer. Reponse to consultee, commentator and public comments on the Appraisal Consultation Documentation (ACD)*. [online] Available at: <http://www.nice.org.uk/guidance/ta250/documents/breast-cancer-advanced-eribulin-response-to-consultee-commentator-and-public-comments-on-the-appraisal-consultation-document-acd2> [Accessed 22 August 2013].

National Institute for Health and Care Excellence, 2012e. *Fingolimod for the treatment of relapsing-remitting multiple sclerosis. Reponse to consultee, commentator and public comments on the Appraisal Consultation Documentation (ACD)*. [online] Available at: <http://www.nice.org.uk/guidance/ta254/documents/multiple-sclerosis-relapsingremitting-fingolimod-response-to-consultee-commentator-and-public-comments-on-the-appraisal-consultation-document-acd2> [Accessed 10 November 2013].

National Institute for Health and Care Excellence, 2012f. *Epilepsy (partial) – retigabine (adjuvant): Patient group, professional group and NHS organisation submission*

---

<sup>38</sup> This list contains additional sources that informed the data analysis but that were not referred to in the thesis text.

*statements*. [online] Available at: <http://www.nice.org.uk/guidance/ta232/documents> [Accessed 25 August 2012].

National Institute for Health and Care Excellence, 2012g. *Hepatitis C (genotype 1) – telaprevir: patient group, professional group and NHS organisation submission statements*. [online] Available at: <http://www.nice.org.uk/guidance/ta252/documents> [Accessed 14 April 2013].