

**The London School of Economics and Political
Science**

Toward Effective Health Technology Regulation

Corinna Sorenson

**A thesis submitted to the Department of Social Policy of the London
School of Economics for the degree of Doctor of Philosophy,
London, February 2015**

Declaration

I certify that the thesis presented herein for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work, other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified). I am responsible for any mistakes.

The copyright of this thesis rests with the author. Quotation from the thesis is permitted, provided that full acknowledgement is made. This thesis may not be reproduced with the prior written consent of the author.

I warrant that this authorisation does not, to the best of my belief, infringe on the rights of any third party.

I declare that this thesis consists of 99,952 words.

Abstract

New health technologies offer both challenges and opportunities. Regulation is one mechanism to help balance the benefits and risks of new health technologies. This thesis examines the extent to which ‘good’ health technology regulation is achieved and the effectiveness of the policy measures regulators (and others) employ to meet such aims. To accomplish these objectives, a conceptual framework of ‘good regulation’ based on the academic and practitioner literatures was developed and its various dimensions considered and explored across eight different studies. Taken together, the studies provide an analysis of the roles, processes, policies, and performance of the regulators responsible for the market authorisation and coverage and reimbursement of pharmaceuticals and medical devices in Europe and the US; the role and use of technology assessment in health technology regulation and its impact on attaining good regulation; and, the factors that impact regulatory performance. The thesis demonstrates that attaining good health technology regulation is complex and challenging, because of the inherent uncertainty regarding the benefits and risks of new technologies, their growing diversity and complexity, the limitations of existing study designs and assessment methods, the increased demands placed on regulators to meet sometimes conflicting objectives, and the underlying political nature of making decisions about public access to and financing of new health technologies. Regulators have made progress on addressing these challenges. However, additional improvements are needed to improve health technology regulatory performance. Like much of health care policy, movement toward achieving the various criteria of good regulation will be incremental, especially considering the often step-wise nature of technological innovation.

Acknowledgements

I would like to extend my sincere thanks to my supervisors, Professor Elias Mossialos and Dr. Adam Oliver, for giving me the opportunity to pursue this research and for their support, guidance, and encouragement. I am immensely grateful for their belief in me throughout the PhD and for allowing me to pursue my (many) interests and priorities along the way. I would also like to thank the following colleagues for their collaboration, support, encouragement, and inspiration at various points throughout the PhD: Kalipso Chalkidou, Michael Drummond, Marin Gemmill-Toyama, Michael Gusmano, Champa Heidbrink, Martin Knapp, Julian Le Grand, Chantel Morel, Irene Papanicolas, Sarah Thomson, Aleksandra Torbica, Yves Verboven, and Grahame Wilkinson. I would particularly like to thank and acknowledge the co-authors of the various studies (Lawton Burns, Kalipso Chalkidou, Michael Drummond, Michael Gusmano, Elias Mossialos, Adam Oliver, and Govin Permanand) and again Lawton Burns for his collaboration on the second study and Karine Chevreul, Isabelle Durand-Zaleski, and Julia Kreis for their collaboration on the fifth study. In addition, I would like to express my gratitude to the experts who participated in interviews for their time and invaluable contributions. Finally, my sincerest thanks to my thesis examiners, Larry Brown and Alistair McGuire, for their astute insights and thoughtful discussion on the thesis and health technology regulation more broadly.

A wise friend told me that developing an academic career is “a marathon, not a race”.... Indeed, and I could not have maintained the vision and fortitude required to complete this phase of the journey without the dedication, understanding, patience, and positive reinforcement of my family and friends. Finally, I would like to acknowledge the many professors and supervisors that have stimulated and encouraged my academic pursuits throughout the years, particularly David Stein, Tamara Ferguson, Peter Jacobsen, Paula Lantz, Kelly Cronin, and Eric Gemmen.

This thesis is dedicated to my parents and my husband, Mark, for their unwavering love and support and for encouraging me to remain steadfast in achieving my dreams and aspirations.

Table of contents

Note on the structure, provenance, peer review, and publication of the thesis	13
Introduction.....	20
Brief overview of the focus of the thesis	20
Toward a framework of ‘good regulation’	21
Why focus on health technology regulation?.....	34
Contributions of the thesis	44
Summary of the thesis methods	56
SECTION I: MARKET AUTHORISATION OF HEALTH TECHNOLOGIES	
Study 1: ‘Good regulation’ and public health protection in the European Union: Evaluating the European Medicines Agency	66
Introduction.....	66
The EMA and public health protection.....	68
Toward a framework of ‘good regulation’	71
Evaluating the EMA	79
Discussion	107
Conclusions.....	113
Study 2: Improving medical device regulation: The United States and Europe in perspective.....	115
Introduction.....	115
Comparative overview of US and European medical device regulation	116
Outstanding challenges in US and European medical device regulation.....	123
Current reforms to improve medical device regulation	128
Additional actions for high performing medical device regulation	134
Concluding remarks	142

SECTION II: HEALTH TECHNOLOGY COVERAGE, REIMBURSEMENT, AND PRICING POLICIES

Study 3: Comparative analysis of pharmaceutical coverage and pricing in Europe: Policy levers and mechanisms and insights for the United States . 144

Introduction.....	144
National policy measures to control pharmaceutical prices and set coverage levels	145
Determining pharmaceutical pricing and coverage	153
Conclusions and insights for the United States.....	166

Study 4: Evolving reimbursement and pricing policies for devices in Europe and the United States and considerations of value 171

Introduction.....	171
Device reimbursement and pricing	172
Policies to improve value-based reimbursement and device pricing.....	180
Conclusion	184

SECTION III: IMPLEMENTATION CHALLENGES ASSOCIATED WITH POLICIES TO IMPROVE HEALTH TECHNOLOGY REGULATION

Study 5: The evolution and impact of health technology assessment in Europe187

Introduction.....	187
Key developments in HTA	188
Evaluations of impact	201
Conclusions.....	206

Study 6: Decision making under uncertainty: An investigation of international coverage with evidence development policies in the context of medical devices 210

Introduction.....	210
Methods.....	212
Results.....	214
Overview of International approaches to CED	214

Key challenges for CED policies	223
Discussion	233
Conclusions.....	239
Study 7: Valuing end of life care: The case of advanced cancer drugs in the United States.....	240
Introduction.....	240
Waging the war on cancer.....	241
The price tag on progress: costs of and access to new cancer drugs	242
Evidence on the value for money of end of life cancer drug treatments	247
The road to value-based funding for and access to cancer drugs: Technical, political and social challenges.....	249
Conclusions.....	253
Study 8: The politics of comparative effectiveness research in the United States: Lessons from recent history.....	259
Introduction.....	259
Setting the policy context.....	260
Past Federal attempts to formalise CER and related approaches.....	263
Lessons learned.....	278
Discussion	281
Conclusion	285
Conclusions of the thesis.....	286
Summary of key thesis finding	286
Policy and practical recommendations to improve health technology regulation	307
Limitations of the thesis.....	312
Further research	315
References	318
Appendix.....	389

List of tables

Table 1: Conceptual framework of ‘good regulation’	32
Table 2: Overview and contribution of each paper	49
Table 3: Advantages and disadvantages of documentary research and expert interview approaches	59
Table 4: Criteria of ‘good regulation’: An evaluative framework.....	77
Table 5: US classifications of medical devices and regulatory requirements for approval and post-marketing surveillance	120
Table 6: European classifications of medical devices and regulatory requirements for approval and post-market surveillance.....	121
Table 7: Overview of existing weaknesses in US medical device regulation, recent reforms, and additional actions for improvement	139
Table 8: Overview of existing weaknesses in European medical device regulation, recent reforms, and additional actions for improvement	140
Table 9: Strategies to control outpatient drug prices and set coverage and reimbursement levels	150
Table 10: Comparative approaches to pharmaceutical pricing and coverage decisions.....	162
Table 11: Measures to enhance evidence-based pharmaceutical coverage, reimbursement, and pricing decisions.....	166
Table 12: Comparative overview of international CED programs	216
Table 13: Summary of past federal attempts to formalise CER and related approaches.....	275

List of figures

Figure 1: Regulators and decisions involved in market access for new technologies	35
--	----

List of boxes

Box 1: Examples of important new medicines with public health benefits approved by the EMA.....	100
Box 2: Example of a device HTA in the UK.....	174
Box 3: Select CED schemes in Canada, UK, and the US.....	219

List of abbreviations

ACA	Affordable Care Act
ACC	American College of Cardiology
ADR	adverse drug reaction
AMA	American Medical Association
AHRQ	Agency for Healthcare Research and Quality
AMNOG	<i>Neuordnung des Arzneimittelmarktes</i>
ASCO	American Society of Clinical Oncology
AWR	approval with research
CCN	Cardiac Care Network
CE	<i>Conformité Européenne</i>
CED	coverage with evidence development
CEESP	Commission for Economic Evaluation and Public Health
CEPS	Healthcare Products Pricing Committee
CER	comparative effectiveness research
CHCT	Council of Health Care Technology
CMS	Centers for Medicare and Medicaid Services
CHCT	Council on Health Care Technology
CHMP	Committee for Medicinal Products
CNEDiMTS	National Committee for the Evaluation of Medical Devices and Health Technologies
COMP	Committee for Orphan Medicinal Products
CPMP	Committee for Proprietary Medicinal Products
CRS	Congressional Research Service
CTA	64-slice CT angioplasty
CVZ	Health Insurance Board
DAHTA	Germany Agency of Health Technology Assessment
DERP	Drug Effectiveness Review Project
DES	drug-eluting stents
DHHS	Department of Health and Human Services
DIMDI	German Institute for Medical Documentation and Information
DRG	diagnosis-related group
EAC	External Assessment Centres

EBM	evidence-based medicine
EEA	European Economic Areas
EMA	European Medicines Agency
EP	extracorporeal photopheresis
EPAR	European Public Assessment Report
EU	European Union
EUDAMED	European Databank on Medical Devices
EUnetHTA	European Network on Health Technology Assessment
EVAR	Endovascular abdominal aortic aneurysm repair
FDA	(US) Food and Drug Administration
GAO	(US) Government Accountability Office
GBA	Federal Joint Committee
GCP	good clinical practice
GMP	good manufacturing practice
HAS	French national Authority for Health
HCFA	Health Care Financing Administration
HMPC	Committee on Herbal Medicinal Products
HTA	health technology assessment
ICES	Institute of Clinical Evaluative Services
ICDs	implantable cardioverter defibrillators
IHC	International Conference on Harmonisation
IOM	Institute of Medicine
IQWiG	Institute for Quality and Efficiency in Healthcare
LVRs	lung volume reduction surgery
MDCG	Medical Device Coordination Group
MDEpiNET	Medical Device Epidemiology Network Initiative
MDR	medical device reporting
MDUFMA	Medical Device User Fee and Modernization Act (MDUFMA)
MedSun	Medical Surveillance Network
MHRA	Medicines and Healthcare Products Regulatory Agency
MRI	magnetic resonance imaging
MTEP	Medical Technology Evaluation Programme
NASS	North American Spine Society
NCD	National Coverage Determination

NCI	National Cancer Institute
NCHCT	National Center for Health Care Technology
NCHSR	National Center for Health Services Research
NICE	National Institute for Health and Care Excellence
NIH	National Institutes of Health
NIHR	National Institute for Health Research
NHS	National Health Service
NOI	notice of intent
NZA	Dutch Healthcare Authority
OECD	Organisation for Economic Co-operation and Development
OHTAC	Ontario Health Technology Advisory Committee (OHTAC)
OIR	only in research
OTA	Office of Technology Assessment
PATH	Program for the Assessment of Technologies in Health
PCORI	Patient-Centered Outcomes Research Institute
PDCO	Paediatric Committee for Medicinal Products
PET	positron emission tomography
PIP	PolyImplant Protheses
PMA	pre-market authorisation
PRO	patient-report outcomes
QALY	quality-adjusted life years
PRAC	Pharmacovigilance Risk Assessment Committee
PSUR	periodic safety update reports
PTA	percutaneous transluminal angioplasty and stenting
RCT	randomised controlled trial
SAMPRIS	stenting and aggressive medical management for preventing recurrent stroke in intracranial
SCARR	Swedish Coronary Angiography and Angioplasty Registry
SEA	Single European Act
SME	small-to-medium enterprise
SmPC	Summary of Product Characteristics
STA	single technology appraisal
TAVI	transcatheter aortic value implantation
TAVR	transcatheter aortic value replacement

THETA	Toronto Health Economic and Technology Assessment Collaboration
TLV	Pharmaceutical Benefits Board
UDI	unique device identifier
US	United States (of America)
UK	United Kingdom
VBID	value-based insurance design
VBP	value-based pricing
VHA	Veterans Health Administration
ZonMW	Netherlands Organisation for Health Research and Development

Note on the structure, provenance, peer review, and publication of the thesis

Structure of the thesis

This thesis follows the publishable paper format, in which a series of papers are submitted as a thesis. The papers must be thematically linked and tied together with an introduction and a conclusion. The introduction discusses the focus of the thesis, its rationale, the contributions of the overall thesis and individual papers, and the overall methodological approach used in the thesis (and across the various papers). The complete papers follow. The thesis then concludes with a summary of the key conclusions across the body of work presented herein, references, and other supporting material (e.g. appendices).

Provenance, peer review, and publication of thesis papers

Study 1

The first study of the thesis is primarily the work of the PhD author (CS). It has been submitted for publication in March 2014 at *Public Administration* and is currently under review.

CS, Elias Mossialos (EM; Brian Abel-Smith Professor of Health Policy, Department of Social Policy, London School of Economics and Political Science), and Govin Permanand (GP; Programme Manager of the Health Evidence Network at the World Health Organization Regional Office for Europe, Denmark and Visiting Research Fellow, LSE Health, London School of Economics and Political Science) devised the paper. CS drafted the paper with contributions from GP. GP and EM reviewed and commented on drafts of the paper. CS finalised the paper for journal submission. In total, CS contributed 80% of the work.

Study 2

The second study of the thesis is primarily the work of the PhD author (CS). In 2014, the paper was published as:

Sorenson, C. and M. Drummond (2014). Improving medical device regulation: The United States and Europe in perspective. *Milbank Quarterly* 92(1): 112-148.

Prior to acceptance, it was subject to double-blind peer review by three referees.

CS devised the paper, reviewed the literature and secondary data sources, and drafted the paper. Michael Drummond (MD; Professor of Health Economics, University of York) commented on drafts of the paper. CS prepared the final paper for journal submission and addressed reviewer comments prior to final submission and acceptance. The paper also benefited from input and comments from Lawton Burns (LB; James Joo-Jin Kim Professor of Health Care Management, Wharton School, University of Pennsylvania). The paper was part of a larger study conducted for The Commonwealth Fund (New York, United States) on medical device policy and was therefore reviewed and commented on by Fund staff, principally Robin Osborn (Vice President and Director, International Program in Health Policy and Innovation). In total, CS contributed 95% of the work.

Study 3

The third study of the thesis is primarily the work of the PhD author (CS). Parts of the paper served as the basis for a paper published by The Commonwealth Fund:

Sorenson, C. (2010). *Use of Comparative Effectiveness Research in Drug Coverage and Pricing Decisions: A Six-Country Comparison*. New York: The Commonwealth Fund.

CS and EM devised the paper. CS conducted the literature review and drafted the paper. EM reviewed and commented on drafts of the paper. CS finalised the paper. In total, CS contributed 95% of the work.

Parts of the paper drew upon earlier work by CS, particularly material specific to drug reviews. However, this study substantially updates the previous work to address the quickly evolving nature of this policy area. The previous work was published as a book in 2008:

Sorenson, C., M. Drummond, and P. Kanavos (2008). *Ensuring value for money in health care: the role of health technology assessment in the European Union*. Observatory Studies Series, 11. Copenhagen, Denmark: World Health Organization, Regional Office for Europe.

The book is based on a comprehensive literature review and interviews of select experts in six EU member states (Finland, France, Germany, the Netherlands, Sweden, and the United Kingdom). The study underpinning the book was a year-long project under the title ‘*Financing Sustainable Health Care in Europe*’, which was endorsed by the Luxembourg Ministry of Health, Sitra, and the Finnish Innovation Fund, and funded by Pfizer.

CS and EM devised the book. CS reviewed the literature and other secondary data sources, interviewed experts, and wrote the book. MD reviewed and commented on the book, with some input from Panos Kanavos (PK; Reader, Department of Social Policy, London School of Economics and Political Science) and EM on the final draft. Drafts of the book were also peer reviewed by David Taylor (Professor of Pharmaceutical and Public Health Policy, University College London) and Frans Rutten (Professor of Health Economics, Institute of Health Policy and Management, Erasmus University Rotterdam) and Willy Palm (Dissemination Development Officer, European Observatory on Health Systems and Policies). CS addressed all reviewer comments and readied the book for publication. In total, CS contributed 95% of the work.

Study 4

The fourth study of the thesis is primarily the work of the PhD author (CS). In 2013, the paper was published as:

Sorenson, C., M. Drummond, and L.R. Burns (2013). Evolving reimbursement and pricing policies for devices in Europe and the United States should encourage greater value. *Health Affairs* 32(4): 788-796.

Prior to publication in the journal, it was subject to double-blind peer review by two referees. It was also critically reviewed by the senior editorial staff at Health Affairs.

CS devised the paper, reviewed the literature and other secondary data sources, and drafted the paper. MD and LB reviewed and commented on drafts of the paper. CS addressed the reviewer comments and prepared the final paper for journal submission. The paper was part of a larger report prepared for The Commonwealth Fund (New York, United States) on medical device policy and was therefore reviewed and commented on by Fund staff, principally Robin Osborn (Vice President and Director, International Program in Health Policy and Innovation). In total, CS contributed 90% of the work.

Study 5

The fifth study of the thesis is primarily the work of the PhD author (CS). In 2012, the paper was published as:

Sorenson, C. and K. Chalkidou (2012). Reflections on the evolution of health technology assessment in Europe. *Health Economics, Policy and Law* 7(4): 25-45.

Prior to publication in the journal, it was subject to double-blind peer review by two referees. It was also critically reviewed by senior policy analysts at the King's Fund, a London-based think tank, who organised the special issue of the journal, which examined developments in European health policy over the last 10 years. In

particular, Anna Dixon and Emmi Poteliakhoff provided valuable comments on early drafts of the paper. Furthermore, Julia Kreis (JK), Karine Chevreul (KC), and Isabelle Durand-Zaleski (IDZ) provided helpful information on key developments in health technology assessment in Germany (JK) and France (KC and IDZ). An earlier draft of the paper was also presented and critically discussed by Carols Gouveia and 30-40 members of the European Health Policy Group (EHPG) during the 10th anniversary EHPG meeting held in London in September 2010.

CS devised the paper, reviewed the literature and other secondary data sources, and wrote the paper. Kalipso Chalkidou (KC; Director, NICE International and Visiting Faculty, Berman Institute of Bioethics, Johns Hopkins University) reviewed and commented on drafts of the paper. CS addressed all reviewer comments and finalised the paper for journal submission. In total, CS contributed 95% of the work.

The paper draws upon and updates earlier work by CS. This work was published as a book in 2008:

Sorenson, C., M. Drummond, and P. Kanavos (2008). *Ensuring Value for Money in Health Care: the Role of Health Technology Assessment in the European Union*. Observatory Studies Series, 11. Copenhagen, Denmark: World Health Organization, Regional Office for Europe.

See study 3 above for further details on the methods underpinning the book, author provenance, and details regarding peer review and publication.

Study 6

The sixth study of the thesis is principally the work of the PhD author (CS). It constitutes a longer paper that has been reformatted into two separate papers submitted for publication in December 2013 and January 2014. In particular, the larger paper (which is presented in the thesis) was fashioned into one paper comparing coverage with evidence development (CED) policies across seven countries and another paper examining the application of CED specifically to medical devices. The former paper was submitted to *Value in Health* and the latter

was submitted to *Health Policy*; the first paper is under review, while the second has been provisionally accepted.

The paper draws on a literature review and expert (health technology assessment bodies/policy makers, industry representatives, and academics/policy analysts) interviews. The experts represented seven countries in North America and Europe.

CS devised the paper, reviewed the literature and other secondary data sources, developed the interview instruments, conducted the interviews, analysed the interview data, and wrote the paper. MD reviewed and commented on the interview instruments and drafts of the paper. Six medical device coverage and reimbursement experts throughout Europe also reviewed the paper. In addition, the study was also presented at two European conferences, which provided the opportunity to verify the accuracy and relevancy of the findings. The presentation won Best Podium Presentation at one of the conferences, which is an award based on the relevancy of the research, methods, meeting the study aims, and clarity of presentation. CS finalised the papers for journal submission. In total, CS contributed 90% of the work.

Study 7

The seventh study of the thesis is solely the work of the PhD author (CS). In 2012, it was published as:

Sorenson, C. (2012). Valuing end-of-life care in the United States: The case of new cancer drugs. *Health Economics, Policy and Law* 7(4): 411-430.

Prior to publication, the paper underwent double-blinded peer review by one referee. The paper was also presented and discussed amongst the LSE-Columbia Health Policy Group in December 2010. MD and KC also reviewed and commented on early drafts of the paper.

CS devised the paper, reviewed the literature and other secondary data sources, and wrote the paper. CS addressed all reviewer comments and prepared the final

paper for journal submission. In total, CS contributed 100% of the work (sole-authored).

Study 8

The eighth and final study of the thesis is primarily the work of the PhD author (CS). The paper was published in 2013 as follows:

Sorenson, C., M. Gusmano, and A. Oliver (2013). The politics of comparative effectiveness research: Lessons from recent history. *Journal of Health Politics, Policy and Law*; published ahead of print November 5, 2013, doi:10.1215/03616878-2396181.

Prior to publication in the journal, the paper was subject to double-blind peer review by four referees.

CS, Michael Gusmano (MG; Research Scholar, The Hastings Center, and adjunct faculty member at Columbia and Yale Universities), and Adam Oliver (AO; Reader, Department of Social Policy, London School of Economics and Political Science) devised the paper. CS reviewed the literature and other secondary data sources and wrote the paper. MG and AO reviewed and commented on drafts of the paper. CS addressed all reviewer comments and prepared the final paper for journal submission. In total, CS contributed 90% of the work.

Introduction

Brief overview of the focus of the thesis

New health technologies offer both challenges and opportunities. They promise hoped-for improvements in health, reduced spending on health from public authorities, and economic development. However, they can also come with risks to individual and population health and public budgets if technologies are adopted and financed that are unsafe, ineffective, or used inappropriately. Regulation is one mechanism to help balance the benefits and risks of new health technologies. Regulatory institutions, policies, and processes have been developed by governments to meet these objectives, namely to authorise technologies for use on the market and to determine the terms of their coverage, reimbursement, and pricing. In this context, regulators aim to ensure that the risk-benefit balance from the safety, efficacy, and quality of the technologies they review is acceptable in the context that they are to be used, and that available budgets are allocated effectively to allow use of such technologies in clinical practice. In practice, however, the regulation of health technologies has inherent limitations and challenges. Industry often maintains that the regulatory process is unpredictable and protracted, thereby thwarting innovation and timely market access of their products; regulators frequently face political pressures, expert-citizen contestation, and stakeholder resistance, leading to problems with compliance or hastened approval processes that may introduce later safety risks or actualised injury; payers are sometimes faced with making coverage and reimbursement decisions based on poor or limited evidence of value; and patient groups and the public frequently decry any restrictions of access to beneficial new technologies.

To that end, this thesis examines the extent to which ‘good health technology’¹ regulation’ is achieved and the effectiveness of the policy measures regulators (and others) employ to meet the aims or criteria of ‘good regulation’. Indeed, good regulation is achieved through a set of tools, activities, and discourses through which different regulators (and involved governments, institutions, and other actors)

¹ While health technology can include drugs, devices, biologics, medical and surgical procedures, support systems, and organisational and managerial systems, the thesis focuses on drugs and devices.

address their policy objectives and reform priorities. To that end, regulators have increasingly relied upon evidence-based approaches to regulation, namely the use of some form of technology assessment to ascertain or substantiate a technology's safety, efficacy, and comparative effectiveness and cost-effectiveness. Such processes typically engage different stakeholders and experts in regulation and, overall, aim to enhance the transparency, accountability, impartiality, and effectiveness of the regulatory process.

Toward a framework of 'good regulation'

If good health technology regulation is to be pursued, it is important to first elucidate a conceptual framework of how 'good regulation' is defined and measured. Yet, developing such a framework raises a number of challenges. First and fundamentally, deriving at a consensus on the definition of 'regulation' is difficult. As Baldwin et al. (1998: pg. 2) note "there is no single agreed meaning of the term [regulation], but rather a variety of definitions in usage that are not reducible to some platonic essence or single concept". The concept itself is somewhat ambiguous; it can be used in both a broad and narrow sense and can encapsulate hard, soft, and self-regulatory approaches. Nonetheless, Baldwin et al. (2012) suggest thinking of regulation in different ways: 1) as a specific set of commands, 2) deliberate state intervention (in the economy or the private sphere), and 3) all forms of social or economic influence (including non-intentional and non-state mechanisms). This thesis is primarily concerned with the first conceptualisation, viewing regulation as a) goal formation, rule-making, and standard setting; b) monitoring, information-gathering, scrutiny, inspection, evaluation, and audit; and, c) enforcement, behaviour-modification, and the application of rewards and sanctions (Hood, Rothstein and Baldwin 2001). These functions may be carried out by a single organisation or delegated separately to specialised agencies. The concept of regulation is often considered an activity that restricts behaviour and prevents certain "undesirables". However, a broader view, which is one this thesis adopts, is that the influence of regulation may also be enabling or facilitative (Baldwin et al. 2012).

Such definitional issues relate to a second challenge. Radaelli and De Francesco (2007: 83) highlight that "...the concept of regulatory quality is prismatic",

and they rightly question whether there is sufficient agreement on what constitutes good regulation, especially across different institutional settings. Indeed, conceptions of quality are likely to vary according to audience constituency, market position, or even discipline (Weatherill 2007). This is often due to divergent weightings of the various criteria or dimensions underpinning good regulation.

Consequently, defining (and indeed measuring) good regulation in all cases is not possible. Nonetheless, it remains important to develop frameworks that can transcend multiple jurisdictions, sectors, institutional settings, and affected actors. To meet this aim, both scholars and practitioners have produced a notable body of literature examining regulation at various levels, offering a range of concepts, ideas, and understandings. While a review of the entire literature is beyond the scope of this thesis, major themes can be identified from several streams of relevant research and drawn upon to develop a framework of good regulation.

Academic approaches² to good regulation tend to be theoretical and less prescriptive in relation to the pursuit of specific outcomes than practical or empirical conceptions. The economics literature is focused largely on (economic) efficiency³ and high productivity (via encouraging investment and innovation) as the primary indicators of regulatory performance, especially when examining public service sectors (e.g. utilities, transportation) (Cubbin and Stern, 2004; den Hertog 2010; Fink et al. 2003; Peltzman 1989). Alternatively, the literature from political theorists and empirical-oriented political scientists emphasise the importance of achieving and furthering accountability, transparency, legitimacy, and other procedural-oriented values (e.g. responsibility, control, openness, predictability, and responsiveness) (Black 2008; Johannsen et al. 2004; Majone 2001; Mulgan 2000; Powell and DiMaggio 1991; Scharpf 1999; Stern 1997).

Taking into consideration both economic and political science perspectives, Baldwin, Cave, and Lodge (2012) put forth the most comprehensive framework based on five criteria or ‘tests’ of good regulation and is oriented toward “...those

² The academic literature on good regulation covers a range of disciplines. I focus on the economic and political science literatures principally; other disciplines of inquiry include philosophy, socio-legal studies, and organisational studies.

³ Namely, that regulation is good if it is efficient in the sense that it maximises wealth.

arguments that have a general currency when regulatory arrangements and performance are discussed in the public domain” (26). The five criteria include mandate, accountability, due process, expertise and impartiality, and effectiveness. These criteria or tests aim to transcend the biases of using efficiency (or any other one indicator) as a single measuring rod or justification for regulation. Moreover, they are applicable to both the instruments of regulation and the regulatory authorities that execute them.

Majone (1996: 300) offers a similar view to Baldwin et al. (2012), where regulatory agencies require a combination of ‘control mechanisms’ to ensure their legitimacy, which he identifies as: “...clear and limited statutory objectives to provide unambiguous performance standards; reason-giving and transparency requirements to facilitate judicial review and public participation; due process provisions to ensure fairness among the inevitable winners and losers from regulatory decisions; and, professionalism to withstand external interference and reduce the risk of an arbitrary use of agency discretion”.

In parallel with academics, international bodies and intergovernmental organisations concerned with regulation at various levels, in particular the Organisation for Economic Co-operation and Development (OECD) and the World Bank, have published practical guidelines, typically employing scorecard approaches on regulatory reform, what constitutes good regulation, and what impact it may have in practice⁴. Much of this work is concerned with economic regulation, often focusing on individual sectors or on issues associated with deregulation. As these guidelines are generally underpinned by an interest in promoting good regulation, by way of a better economic environment, several common features can be identified. Such elements include that regulation 1) have a strong legal basis (regulators must be independent), 2) be clear and feasible to implement, 3) bring a net benefit, and 4) is efficient. These dimensions are echoed in various national level guidelines, with notable examples including the work of the Australian Office of Regulation Review (AORR 1998), the Council of Australian Governments (2007), the United

⁴ See Jaramillo 2013; Kaufmann et al. 1999, 2002, 2004, 2005, 2009; and, OECD 1997, 2005.

Kingdom's Better Regulation Task Force (BRTF 2000)⁵, and the External Advisory Committee on Smart Regulation to the Canadian Government (CEACSM 2004). Again, these are largely economic focused, but in being issued by elected governments, they do take wider social concerns into account.

This is also true of related guidance put forward at the European level. For example, the Mandelkern Report on Better Regulation (COM 2001a) served as a basis for drawing the Better Regulation Policy in the European Union (EU). In parallel, the 2001 White Paper (COM 2001b) on European governance⁶ outlined five principles – openness, participation, accountability, effectiveness and coherence – aimed at engendering and maintaining trust in the way the EU governs. These principles are particularly important for those agencies whose remit carries direct social policy impacts. More recent publications of the European Commission focus on the importance of responsiveness in good regulation – meaning that regulatory policies and tools undergo frequent modification in light of experience and evolving needs to ensure that they are “fit for purpose” and moderate any undue administrative burdens (COM 2010a; COM 2010b; COM 2012). Such objectives are reflective in the academic literature on “responsive regulation” (Braithwaite 2006; Neilsen and Parker 2009; Ojo 2009), “risk-based regulation” (Baldwin and Black 2007), and “really responsive regulation” (Baldwin and Black 2007; Black 2008).

The conceptual framework of good regulation employed in this thesis draws principally upon the work of Baldwin and colleagues, but is also informed and complemented by elements from the aforementioned practitioner-oriented guidelines. While the former represents objective dimensions for assessing regulatory performance, it is largely indicative. As such, the practitioner criteria are therefore considered complementary to Baldwin et al.'s more overarching conceptual principles. Together, they help form a framework that encompasses both theoretical and empirical considerations of good regulation. In particular, the framework

⁵ See also the Regulatory Impact Unit's work on the effective undertaking of regulatory impact assessments (RIU, 2003).

⁶ While distinct concepts, good regulation can be seen as an element or exercise of good governance, with similar criteria and principles applicable to both constructs. The OECD (2001) and Kaufmann et al. (2002) have both set out criteria of good governance, and then specifically linked them to regulatory agencies. Moreover, as much work on European integration considers regulation as the hallmark of the EU in terms of a *sui generis* form of governance (see Eberlein and Kerwer 2002; Majone 1996), the two concepts can be linked together in this context.

encompasses the five criteria outlined by Baldwin and colleagues (mandate, accountability, due process, expertise and impartiality, and effectiveness), in addition to one other criterion commonly used by practitioners – cost-efficiency. In attempts to remain sensitive to the economic orientation of health technology regulation, the framework adopts the differentiation made by the Canadian government between effectiveness and cost-efficiency. The framework therefore separates regulatory efficiency from regulatory effectiveness, with the latter focused on the delivery of policy results as opposed to questions of cost and allocative efficiencies. Outlined in Table 1, the thesis therefore defines and evaluates good regulation across the following criteria:

Criteria 1: *Mandate*

Regulators are authorised or mandated to assume certain responsibilities or functions, typically by a relevant legislative authority (e.g. Parliament, Congress). Such mandates normally presume a public interest outcome to be served, and success requires that the regulator follows and achieves its particular mandate(s), which allows for claims of public support and legitimacy (Baldwin and Cave, 1999; Baldwin et al. 2012; Kaufmann et al. 1999, 2002, 2009). They also suppose that the regulator and associated actions are proportional⁷ and consistent with other national or international regulations, law, or policies (Australian Office of Regulation Review 1998; Canadian External Advisory Committee 2004; European Commission 2001a; Kaufmann et al. 1999; UK Better Regulation Task Force 2000). Furthermore, a regulator's mandate should be clear to ensure it is enforceable and can be assessed for performance against its stated objectives (Australian Office of Regulatory Review 1998; OECD 1997, 2003).

In many cases, regulators of health technologies can be categorised as independent regulatory agencies, which are formally independent from direct (ex-ante and ex-post) political control. Legitimization of the agency and its mandate is normally based on a large array of “non-democratic” justifications, but primarily the

⁷ The proportionality principle states that the means used to achieve certain ends must be necessary and least burdensome, hence, the minimum necessary to reach a certain goal.

need for insulation from day-to-day politics and the exercise of specific policy competencies (e.g. technical expertise) (Majone 1996).

Criteria 2: *Accountability*

Further claims for support or legitimacy can be made when the second criterion –accountability– is achieved. In most cases, health technology regulatory bodies are non-elected, which means they are not democratically accountable in the traditional sense of being politically responsive to citizens by way of a chain of political delegation (Maggetti 2010). The question of ensuring regulatory bodies are accountable is indeed a significant concern (Finders 2004; Hood and Scott 2000). Accountability is an expanding terms that means, in its core sense, to be called to account for one's actions, hence presupposing the existence of "external scrutiny" and ability to justify decisions via sufficient reason-giving (Baldwin et al. 2012; Castiglione 2006; European Commission 2001a; Mulgan 2000; UK Better Regulation Task Force 2000). Accountability may encompass different meanings: answerability, responsibility, control, responsiveness, openness, and dialogue with citizens (Baldwin et al. 2012; Mulgan 2000). Bovens (2007) underlines the fact that accountability should be conceptualised as a social relation between the regulator and its "accountability form", which can be an individual actor, or a collective form, such as parliament, government, or a stakeholder group(s).

Therefore, the link between accountability and legitimacy is often conceived in procedural terms. Stakeholders, even if they disagree with a regulatory decision, should accept it as legitimate and justifiable if it was made in a way considered fair and appropriate, namely if it originated from an open and inclusive political process, ideally based on openness, transparency, equal access, and deliberation. Thus, one of the principal objectives of accountability is to establish and maintain public trust in and support of a regulator's mandate, actions, and outputs. This involves the application of a number of measures, such as the development of standards or guidelines for production and service delivery; the presence of interest groups, users, and other stakeholders in overall governance and processes; the employment of performance surveys or evaluations of internal and external review; the availability of public reports about regulators' performance (e.g. annual report); and, more

generally, the improvement or adjustment of regulator's policies and processes to answer, explain, and justify their actions (Lodge 2004; Majone 1997; Scott 2000).

Criteria 3: *Due Process*

The third criterion, due process, largely relates to the decision making and implementation phase of public policy making. Here, the focus is on transparency, fairness, and consistency of treatment, as well as the levels of participation regulators afford the public, consumers, and other affected parties. Due process effectively ensures proper democratic influence over regulation, thereby, in concert with accountability, exercising a legitimising effect and securing public trust. In practice, this involves ensuring a reliable and open flow of information available to affected actors and the public that allows sufficient understanding of what decisions are made and who makes them; the processes and criteria for arriving at said decisions, including any dissenting views; changes in governance, policies, and processes; and, appropriately detailed and reliable information about the activities, achievements, and failures of the regulatory entity. In addition, due process requires availing opportunities for involvement of a wide range of stakeholders in various levels of regulation ('regulatory discretion') and some degree of 'substantive equality', where there are consistencies in regulation, associated processes, and opportunities for stakeholder engagement (Knill and Lenschow 2003).

Criteria 4: *Expertise and Impartiality*

The fourth criterion relates to the fact that regulators are expected to be impartial and have sufficient expertise to exercise judgement in a way acceptable to the public (Thatcher 2002). This is particularly true for delegated agencies, where the efficiency of the regulatory process and the credibility of the agency depend on high quality data and cutting-edge expertise (Genoud 2003). Together, expertise and impartiality engender public trust and support, which allows for an agency to exercise discretion in their work (Baldwin et al. 2012). These elements also confer a

level of professionalism to withstand external interference, avoid regulatory capture⁸, and reduce arbitrary use of discretion in decision making (Majone 1996; OECD 1997, 2003). This is particularly important in situations whether the regulator or decision maker is required to consider a range of competing options, opinions, and/or values and arrive at a balanced judgement on incomplete or shifting evidence. To that end, securing a sufficient level of expertise in regulation may also help ensure that decisions are robust to errors.

Criteria 5: *Effectiveness*

The effectiveness criterion can be best understood as whether a regulator delivers intended results or policy objectives (Baldwin and Cave 1999; Baldwin et al. 2012; OECD 1997, 2003). Several of the practical frameworks for good regulation emphasise that achievement of these aims should be based on standards and targets (Australian Office of Regulation Review 1998; Canadian External Advisory Committee 2004; UK Better Regulation Task Force 2000) or ‘goal-based approaches’ (OECD 1997, 2003). Such standards can originate either internally or externally to the regulatory organisation.

The notion of regulatory responsiveness, as previously discussed, relates to the effectiveness criterion in two overarching ways – the first being that the regulator is organised in such a way that adequately allows it to meet current and future challenges and, second, that the regulator can and does alter its procedures to ongoing needs and challenges, where appropriate. On a more specific level, Baldwin and Black (2007) propose that successful responsiveness entails accounting for different values, opinions, and experiences that operate within regulated bodies and the regulated; responding to the constraints and opportunities presented by institutional and external contexts; receptivity to the logics of different regulatory tools and strategies; and, awareness of regulatory performance and adaptiveness to modification.

⁸ Regulatory capture is the result or process by which regulation, in law or application, is consistently or repeatedly directed away from the public interest and toward the interests of the regulated industry, by the intent and action of the industry itself (Carpenter and Moss 2013).

Criteria 6: *Cost-Efficiency*

Regulation is cost-efficient when the output of regulation justifies the cost. In order to meet this end, it is essential to understand the cumulative impact of policies and to avoid duplication and overlap in regulatory activities. In particular, claims to this test involve both productive efficiency and efficient regulatory outcomes. The former reflects whether the mandate is implemented at the least possible level of inputs or costs. The latter encapsulates whether the regulator or regulation under examination leads to results that are efficient⁹, which unlike productive efficiency, is judged with a degree of independence from the mandate itself.

Furthermore, achieving cost-efficiency must be based on first meeting the first four criteria of the framework. Consequently, it is perhaps the most interesting (and complex) criterion from an evaluation point of view, in that it aims to capture the dichotomy underlying most regulatory policies – the tension between the public and private aspects of regulation. In the case of health technology regulation, the tension is premised on opposing interests between consumers (patients) and producers (industry), with the former focused on protecting societal, public health concerns and the latter maintaining that regulation is designed first to serve industry. It is this strain, in particular, that characterises many of the challenges raised in the first four criteria of the framework. There are therefore definitional problems related to the cost-efficiency criterion, in terms of determining which objectives and whose needs are met first. In other words, how to balance and assess economic efficiency versus social objectives?

Therefore, according to the framework, good health technology regulation can be achieved or enhanced through the following mechanisms:

- Clear and appropriate mandates of involved regulators;

⁹ Outcome efficiency can be judged across two measures: allocative efficiency (whether it is possible to redistribute goods to increase the benefits to or welfare of any one consumer without making another consumer worse off) and dynamic efficiency (whether there is encouragement of desirable process and product innovation, and whether the system produces flexible responses to changes to demand).

- Accountable and independent regulators and associated processes, which are subject to external scrutiny and sufficient “reason-giving” for their policies and decisions;
- Fair, open, transparent, and inclusive regulatory processes and policies;
- Sufficient expert involvement to secure public support and trust in decisions, withstand external interference, reduce arbitrary use of discretion, facilitate evidence-based decision making, and protect against errors;
- Responsiveness to the constraints and opportunities presented by institutional environments within which regulators act, as well as the logics of different regulatory tools and strategies and requirements for change;
- Timely regulatory processes that avoid duplication and overlap and where the benefits justify any costs; and,
- Delivery of intended results or policy objectives.

Certainly, there may be other mechanisms that may play a role (again, different stakeholder groups may differ on what constitutes ‘good regulation’), but the framework offers a comprehensive foundation for examining the different dimensions of regulatory performance and quality.

The framework is intended to be of interest to and potentially used by a variety of affected stakeholders to better understand and assess the processes, policies, and performance of regulators. A broad audience aligns with some of key criteria of good regulation, namely opportunities for stakeholder involvement and external evaluation or scrutiny of regulatory performance. Such stakeholder groups include national policy makers, academics or policy analysts, industry, patient organizations, consumer groups, etc. National policy makers, for example, could employ the framework to evaluate how publicly funded regulatory bodies are performing against the criteria of good regulation. This could inform funding priorities, reform policies, new initiatives, communication and interaction with affected parties, and resource allocation decisions, among others. As mentioned, it would also be useful to industry. Similar to other affected stakeholders, industry could employ the framework to understand how effectively regulators are overseeing and making authorisation and coverage and reimbursement decisions on their products.

Accordingly, manufacturers may be better able to identify areas of reform as well as pressure points to more effectively engage with regulators to ensure satisfactory regulatory performance.

Table 1: Conceptual framework of ‘good regulation’

Mandate	(External) Accountability	Due Process	Expertise & Impartiality	Effectiveness	Cost-Efficiency
Authorised legislative mandate to claim public support (1)(2)	Answerable to elected body representing the public (‘democratically responsive’) (1)(2)	Fairness, openness, transparency, inclusion of relevant stakeholders (1)(7)(8)(9)	‘Sufficient’ expertise in order to secure public support for exercise of discretion (1)(2)	Delivering intended results (1)(2)(4) or policy objectives, as based on standards and targets (6)(7)(8)	Costs and savings generated (1)(2) where benefits justify costs (4)
Able to balance and ensure independence and accountability (3)(4)				Promote innovation through incentives and goal-based approaches (4)	Minimise costs and market distortions (4)
Sound legal basis (4) and regulatory backing (authority) (5)	Able to justify decisions and be subject to public scrutiny (5)(7)(8)(9)	Ensure fairness amongst inevitable winners and losers (3)	Professionalism to withstand external interference and reduce arbitrary use of discretion (3)	Better than alternatives (6)	Understanding cumulative impact of policies (risk and problem awareness) and avoiding duplication and overlap (8)
Stated regulatory objectives (5) which are better than alternatives (6)	Subject to adjustment (6) ^a				

Proportionality (7)(9) and necessity (8)	Subject to appraisal by independent bodies (6) or other external scrutiny (6)(9) in order to avoid 'regulatory capture' (4)		Takes account of the cultures and understanding that operate within regulated bodies; responds to constraints and opportunities presented by institutional environments within which regulator acts; responsiveness to the logics of different regulatory tools and strategies; performance awareness and modification; adaption to change (10) ^b	Timelines (8)
Consistency with other (national and international) regulations/law/policies (5)(6)(7) and, in the EU, respecting subsidiarity(9)	'Reason-giving' and transparency to facilitate judicial review and participation (1)(2)(3)		Robust to errors (6)	
Based on verifiable performance criteria (4)(6)		Clear and practical for users (4)	Evidence-based decision making (8) ^c	
Enforceable (6)		Improving internal management and serving stakeholders (2)		

Sources: Authors' compilation based on: (1) Baldwin and Cave 1999; (2) Baldwin et al. 2012; (3) Majone 1996; (4) OECD 1997, 2003; (5) Kaufmann et al. 1999, 2002, 2004, 2005, 2009; (6) Australian Office of Regulation Review 1998; (7) UK Better Regulation Task Force 2000; (8) Canadian External Advisory Committee 2004; (9) Mandelkern Report, European Commission 2001a; (10) Baldwin and Black 2007.

a This is potentially limited by the mandate.

b Some of these issues also related to accountability.

c The Canadian Report considers evidence-based decision making as an element of the 'effectiveness' criterion, but it is considered applicable also to the requirements of 'expertise' and 'cost-efficiency'.

Why focus on health technology regulation?

Health technology regulation has a long and important history in national and international health policy. With the establishment of the Food and Drug Administration (FDA) in the United States (US) and the European Medicines Agency (EMA) in Europe¹⁰, new health technologies are required to undergo review and assessment to ascertain their benefits and risks to public health before being marketed on the health care system. In other words, new technologies must demonstrate that they do more good than harm in a defined group of patients, where benefits and risks are considered in clinical terms (i.e. will it work?) (Figure 1). Most often, this means that each new technology is evaluated on its own merit, not compared to other available treatments, and under controlled conditions (typically a randomised controlled trial, RCT).

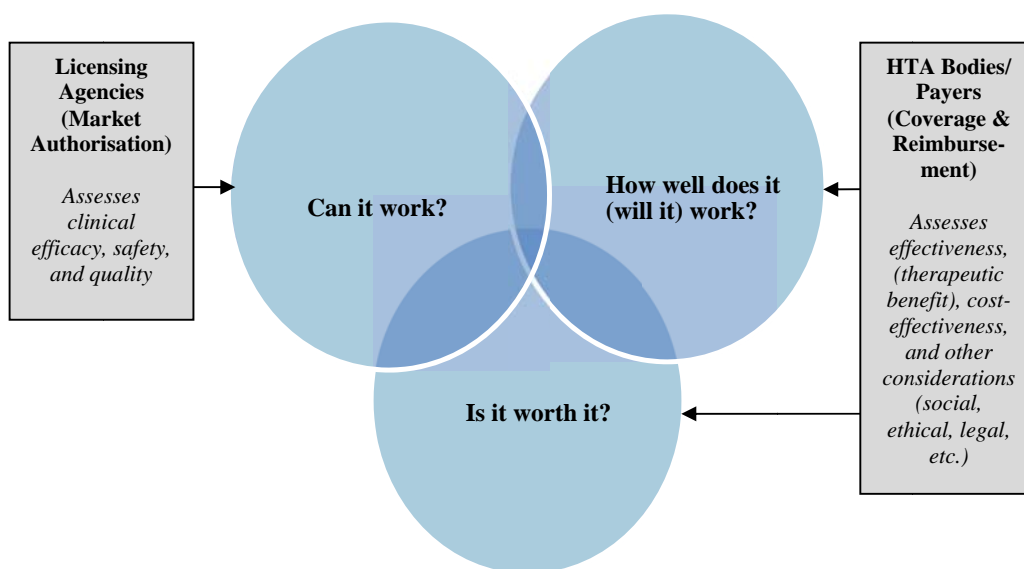
Until the 1990s, market authorisation was the sole hurdle to market access for health technologies. However, over the past two decades or so, national governments and, in particular, entities responsible for the public financing (coverage and reimbursement) of health technologies have instituted similar assessment processes (broadly termed, health technology assessment, or HTA¹¹) to ascertain the health and (sometimes) costs consequences associated with new technology. The overarching goal, in most cases, is to optimise the health outcomes for a population of patients by considering all available treatment options while accounting for budgetary constraints. Therefore, such assessments typically compare the new technology against existing treatment alternatives to address the question of whether a technology provides therapeutic value and, in some circumstances, health system and societal value (Figure 1). Assessments typically involve a broader range of evidentiary approaches (e.g. RCTs, comparative benefit/effectiveness studies, observational studies, health economic modelling). Some government bodies conduct all assessments in-house or avail themselves of (semi-) independent, quasi-regulatory

¹⁰ There are other pharmaceutical licensing bodies, but the FDA and EMA are responsible for approximately 80% of the world pharmaceutical market (McCabe et al. 2008). The FDA also reviews other types of health technologies, namely medical devices.

¹¹ HTA is the broadest term for technology assessments, but similar research has recently been termed 'comparative effectiveness research' or CER, predominately in the US. There are, however, differences. CER typically equates to a comparative assessment of effectiveness only, while HTA includes effectiveness, but also an economic dimension (cost-effectiveness) and may also consider the social, ethical, and legal aspects of health technologies.

entities to elaborate on coverage and reimbursement (and, sometimes, pricing) recommendations or determinations. The National Institute for Health and Care Excellence (NICE) in England and Wales and the Centers for Medicare and Medicaid Services (CMS) in the US are two prominent examples.

Figure 1: *Regulators and decisions involved in market access for new technologies*



The impetus for such regulatory institutions, policies, and processes extends beyond protecting public health and public budgets, while duly facilitating patient access to important new treatments. Greater public awareness of regulatory decisions, a growing consumerist perspective on health care, and increasing demand for well-founded information on health technologies have placed pressure on governments to attain more accountable, transparent, inclusive, and legitimate decision making. The evidence-based approaches employed by many regulators are considered a viable way to meet these procedural objectives (Sorenson et al. 2008a).

The body of academic inquiry on health technology regulation has grown considerably over the last 15 years, particularly with regards to HTA. There are several reasons for this development: increased use of evidenced-based regulation and interest in its potential to further a range of policy goals, particularly during a time of sustained fiscal pressure; greater awareness of the problems with health technology regulation in practice; and, as intimated above, a commitment on the part

of policy makers to improve regulatory processes. These key points are elaborated further below.

Growing interest in and use of evidence-based approaches to health technology regulation

Market authorisation and coverage and reimbursement authorities face uncertain choices when considering the adoption of health care technologies. While consideration of the available evidence on the benefits and risks has traditionally played a role in market authorisation decisions, particularly with regards to pharmaceuticals, this has not always been the case with respect to coverage and reimbursement policy.

Health technology assessment originated in the US in the 1970s in response to mounting concern about the diffusion of costly health technologies and governments' and taxpayers ability and willingness to fund their use. The Office of Technology Assessment (OTA) was established to provide Congress impartial assessments of technologies in medicine and other technology-based sectors that became the basis of many subsequent public policies (Bimber 1986). While the OTA was not involved in coverage and reimbursement policy and was ultimately disbanded in 1995, it served as a model for the creation of similar entities in Europe and elsewhere. Some of these bodies focus on the production of evidence-based reports for use in a broad context (e.g. the Swedish Council on Technology Assessment in Health Care), while others attend to the production of guidance decisions linked to the coverage and reimbursement of health technologies (e.g. NICE in England, the Institute for Quality and Efficiency in Health Care (IGWiG) in Germany, and the Canadian Agency for Drugs and Technologies in Health (CADTH) in Canada). In some countries, such as France, the Netherlands, and Sweden, rather than establishing a stand-alone HTA body to advise or decide on coverage and reimbursement policy, technology assessment processes have been adopted within the purview and operations of an existing government authority.

Since the disappearance of the OTA, the US has experimented with HTA and similar types of policy research (e.g. outcomes research, health services research,

comparative effectiveness research), but its implicit use to inform coverage and reimbursement has been limited, especially at the Federal level. Compared to other countries, the adoption of HTA at the national level, especially to determine access to new technologies, has been most challenging and contentious, in part due to the decentralised public-private health system and the national politics around rationing. The 2010 Affordable Care Act (ACA), however, promulgated the need for evidence on health care interventions and invested substantially in a new institute, the Patient-Centered Outcomes Research Institute (PCORI) to fund and oversee comparative effectiveness research (CER).

Since their inception, these bodies (and technology assessment more broadly) have grown in influence, as a result of an ever-expanding number and breadth of new technology and, in turn, growing evidence demands to ascertain value and safety; expanding health care expenditures and constrained budgets; growing stakeholder demands for information and expedient access to technology; and, the increasingly interconnectedness of national health technology markets. Moreover, it is increasingly the case that new technologies must attain not only market authorisation, but also a positive coverage decision and sufficient level of reimbursement in order to reach patients in a timely way or at all. This has resulted in important consequences for regulators, physicians, patients, and the health system as a whole. First, the decision-making power of coverage and reimbursement bodies has grown, which has effectively replaced some of the decision authority of physicians, as prescribing decisions are becoming more restricted by payer's decisions (Eichler et al. 2010). Second, the ability of new technology, especially those of high expense, to be adopted into practice is increasingly driven by the ability of manufacturers to demonstrate added value to payers. Third, unlike prescribing, but similar to authorisation decisions, coverage and reimbursement decisions are often taken by specialised institutions, expert committees, and increasingly based on a dossier of complex data and sophisticated methodology. Fourth, the dual assessments performed by the two regulators can result in contentious situations (e.g. a technology is approved by a licensing agency on the basis of its safety and efficacy, but is subsequently deemed not reimbursable by payers).

Health technology regulation has therefore become more visible and contentious, as it inevitably brings together public and private interests in a process where there are potentially winners and losers and the perception of outcome is highly contingent on each party's point of view. Indeed, advocates herald the use of technology assessment and resulting evidence in decision making to advance population-based health and promote efficient resource allocation, while critics and sceptics consider such approaches a way to simply restrict access to new technology or displace inherently political choices with technical ones.

As a consequence, academic and policy interest in health technology regulation have risen. Studies have focused on the following areas¹²:

- Regulation of pharmaceuticals (Abraham 1995; Abraham and Lewis, 1998; Abraham and Lewis 1999; Barbui et al. 2011; Bassi et al. 2003; Gardner 1996; Garattini and Bertele 2007; Garattini and Chalmers 2009; Lexchin and Donovan 2010; Mossialos et al. 2006; Regnstrom et al. 2010; Wiktorowicz 2003), and in particular to Europe (Abraham and Lewis 2000; Ernst and Young 2010; Gardner 1996; Mossialos et al. 2004; Permanand and Mossialos 2005; Mossialos and Oliver 2005; Motola et al. 2006; Permanand 2006), the US (Carpenter 2010; Daemmrigh 2004; Kane 1997; Lakdawalla et al. 2009), and across multiple jurisdictions (Banta 1995; Franken et al. 2012; Kanavos 2003; Morgan et al. 2006; Mossialos and Oliver 2005; Vogler et al. 2009).
- Pharmaceutical coverage and reimbursement in Europe (mostly focused on select member states) (Annemans et al. 1997; Barros 2010; Folino-Gallo et al. 2008; Franken et al. 2012; Garattini et al. 2007; Gulasci et al. 2002; Haga and Sverre 2002; Kanavos 2003; Lundkvist 2002; Moise and Docteur 2007; Mossialos et al. 2006; Mossialos and Oliver 2005; Paris and Docteur 2007; Pedersen 2003; Rinta 2001; Rovira and Darba 2001; Stafinski et al. 2011a; Yfantopoulos 2008; Vogler et al. 2009; Vogler et al. 2011), the US (Berndt and Newhouse 2010; Forrest et al. 2005; Neumann et al. 2008), and in other jurisdictions (Lexchin and Mintzes 2008; Paris and Docteur 2006).

¹² List intended to be a comprehensive, but not exhaustive, outline of literature in the field. Discussion of specific literature gaps can also be found in the paper summaries and the individual papers.

- Development and status of health technology assessment and the role or need for evidence-based coverage and reimbursement decision making (Banta 2003; Banta and Oortwijn 2000; Battista and Hodge 1999; Battista 2006; Jonsson 2002; Maynard and McDaid 2003; Perry et al. 1997; Perry and Tharner 1999; Sorenson et al. 2008a; Stevens et al. 2003; Valesco-Garrido et al. 2008);
- Technology assessment in particular jurisdictions, with the United Kingdom (UK) and NICE, in particular, the most heavily researched (Banta et al. 1995; Berg et al. 2004; Bos 2000; Carlsson 2004; Chalkidou and Walley 2010; Culyer 2006; Devlin and Parkin 2004; Drummond and Sorenson 2009; Eisenberg and Zarin 2002; Gerdhardus 2006; Lauslahti et al. 2000; Menon and Topfer 2000; Oliver et al. 2004; Oortwijn et al. 2008; Orvain et al. 2004; Perleth et al. 2009; Rawlins and Culyer 2004; Sorenson et al. 2008b; Stevens and Milner 2004; Sullivan et al. 2009; Woolf and Henshall 2000);
- Health technology assessment institutions and processes across different jurisdictions (Chalkidou et al. 2009; Chinitz 2004; Clement et al. 2009; Draborg et al. 2005; Garcia-Altes et al. 2004; Harris et al. 2001; Lexchin and Mintzes; 2008; Nicod and Kanavos 2012; Sorenson et al. 2008a; Sorenson and Kanavos 2009; Sorenson 2010; Stafinski et al. 2011a; Stafinski et al. 2011b; Oliver et al. 2004; Perry and Tharner 1997; Velasco-Garrido et al. 2008; Zentner et al. 2005) and different therapeutic areas, including orphan conditions (Nicod and Kanavos 2013), cancer (Faden et al. 2009; Mason et al. 2010), and biosimilar drugs (Minghetti et al. 2011);
- Methods and processes of technology assessment (Abelson et al. 2007; Anell 2004; Anell and Persson 2005; Boulenger et al. 2005; Busse et al. 2002; Chalkidou et al. 2007; Claxton et al. 2005; Drummond et al. 2005; Facey et al. 2010; Greenberg et al. 2005; Hutton et al. 2008; Johnson 2009; Lehoux and William-Jones 2007; McGregor and Brophy 2005; Neumann 2004; Neumann et al. 2010; Noorani et al. 2007; Phillips et al. 2006);

- Use of evidence/technology assessment in coverage and reimbursement decisions (Borowski et al. 2007; Carino et al. 2006; Cookson et al. 2001; Drummond et al. 2008; Draborg and Andersen 2005; Hivon et al. 2005; Hutton et al. 2006; Hutton et al. 2007; Luce and Brown 1995; Miller and Pearson 2008; Mohr and Tunis 2010; Oortwijn et al. 2010; Ramsey and Sullivan 2005; Sorenson et al. 2008a; Sorenson et al. 2010; Trueman et al. 2010).

Despite the ever-expanding knowledge base, more research is needed, especially to assess new regulatory developments and evolutions in practice. In particular, there is limited research on market authorisation, especially from the agency perspective and with regards to medical devices. Only a few studies have been published focusing on medical device regulation (Altenstetter 2003; Altenstetter 2008; Altenstetter 2012; Altenstetter and Permanand 2007; Kramar et al. 2012; Kramar et al. 2013). Of those studies focused on pharmaceutical regulation, the majority were conducted five to ten years ago and therefore somewhat outdated.

The evidence on national health technology coverage and reimbursement policy making is more expansive, but, again, existing studies have principally focused on pharmaceuticals. Research on the coverage and reimbursement of devices, including assessment or methodological challenges, has only recently received academic attention (Basu and Hassenplug 2012; Gelijns et al. 2013; Kirisits and Redekop 2013; Schreyogg et al. 2009; Sorenson and Kanavos 2011; Sorenson et al. 2011a; Torbica and Cappellaro 2010). The need for such research is notable, given the significant growth in the medical device industry in recent years and, as a consequence, the increased development and availability of sophisticated, costly devices.

Awareness of challenges associated with health technology regulation in practice

Given the diversity of available health technologies and the complexity of regulating them, there are inherent challenges to regulation in practice. The limitations associated with health technology regulation have been noted in the literature. Key challenges relate to¹⁰:

- Timeliness of regulatory decisions and access to technology (Basu and Hassenplug 2012; Carpenter et al. 2008; Mason and Drummond 2009; Munos 2009); methodological and evidence hurdles (Campillo-Artero 2013; Cookson and Hutton 2003; Drummond 2004; Drummond and Sculpher 2005; Drummond et al. 2009; Drummond 2013; Eichler et al. 2010; Eichler et al. 2011; Naci et al. 2012; Oliver and Sorenson 2009; Sorenson et al. 2011a; Sorenson et al. 2011b);
- Ensuring post-market technology safety, effectiveness, and cost-effectiveness (Abraham and Davis 2005; Furberg et al. 2006; Resnic and Normand 2012);
- Securing public accountability, transparency, and legitimacy (Permanand 2006) and stakeholder involvement (Abelson et al. 2007; Facey et al. 2010; Facey et al. 2011; Gauvin et al. 2010; Milewa and Barry 2005; Milewa 2006);
- Social and political concerns (Avorn 2009; Brown 1991; Carpenter 2006; Cookson and Maynard 2000; Gelijns et al. 2005; Gerber et al. 2010; Gusmano and Gray 2010; Iglehart 2010; Lehoux and Blume 2000; Manchikanti et al. 2010; Oberlander et al. 2001; Permanand 2006; Wilensky 2009);
- Defining and judging value (Valesco-Garrido et al. 2008; Gelijns et al. 2013; Goldman et al. 2010; Hofman 2008; Kennedy 2009; Littlejohns et al. 2012); and,
- Impacts of technology assessment (Drummond and Weatherly 2000; Jacob and McGregor 1997; Sigmund and Kristensen 2002; Oliver et al. 2004; Sorenson et al. 2008b).

Many of the challenges relate particularly to technology assessment and its use in practice to inform decision making, particularly coverage and reimbursement. However, most studies have focused on a specific issue (e.g. stakeholder involvement, economic evaluation) and not on a broader set of issues (and the dynamics between them). In relation, the policies or reforms introduced by regulators and other actors to address said challenges deserves examination.

Commitment of policy makers to improve regulatory processes

Policy makers have increasingly recognised that health technology (and their regulation) forms an integral part of any truly effective modern health care system, and that it plays an important role in ensuring a healthy and productive society. Significant resources are dedicated to funding the activities of regulatory agencies, and various advisory committees have been created to advise policy makers on key issues related to the regulation of health technologies, particularly pharmaceuticals. In addition, there is considerable interest among policy leaders in Europe, particularly with regards to HTA, to enhance cross-border collaboration between involved bodies and identify “core” assessment standards to be shared and exercised internationally, in attempts to make regulation more efficient and predictable (Kristensen et al. 2009a; Kristensen et al. 2009b; Sorenson et al. 2008b). The European Parliament’s recent directive on patient rights and cross-border health care supported more formalised cooperation between national HTA bodies through the European Network of Health Technology Assessment (EUnetHTA) (European Parliament 2011). Similar aims have been sought within the pharmaceutical market authorisation arena with the International Conference on Harmonisation (ICH).

There is also growing interest in understanding and addressing the entire regulatory continuum for health technologies (i.e. the regulatory life-cycle). Historically, market authorisation and coverage and reimbursement processes have been perceived as dissimilar and separate. Academics in the field have generally focused on some aspect(s) of market authorisation or coverage and reimbursement, but rarely together. There has been some work in this area in recent years (Breckenridge 2010; Eichler et al. 2010; Fronsdaal et al. 2012; Henshall et al. 2011; Henshall et al. 2013). Research on the topic has aligned with policy makers’ interest

in ensuring better synergies between market authorisation and coverage and reimbursement. In Europe, for example, policy makers (particularly DG Sanco and DG Enterprise) have discussed the idea of a Europe-wide relative efficacy and/or relative effectiveness assessment for new pharmaceuticals (Kleijnen et al. 2011; Eichler et al. 2010). In the US, the FDA and CMS have started a pilot ‘parallel-review’ programme for medical devices (Messner and Tunis 2012). Interest in this topic amongst policy makers (and academics) is attributable to a number of factors, including a commitment to ensuring patient access to beneficial and appropriate interventions, enhancing the efficiency of the development of new products and their regulation, strengthening the evidence base for decision making, and facilitating the transparency of such decisions and their rationale to the general public.

Contributions of the thesis

Against the aforementioned backdrop, it is therefore an opportune moment to investigate health technology regulatory performance and the measures used to ensure good regulation. This thesis accomplishes these objectives by addressing the following main and sub-questions:

1. To what extent is ‘good health technology regulation’ being achieved?

1a. Does regulatory performance differ between systems for drugs and devices?

1b. Are any of the criteria outlined in the framework more important to attaining good regulation?

2. What factors, if any, inhibit or facilitate meeting the various criteria of good regulation?

3. Have evidence-based approaches to health technology regulation aided regulators ability to achieve and maintain good regulation?

3a. Is the impact of such approaches different for pre- (market authorisation) and post-market (coverage and reimbursement) regulation?

3b. Are they “fit for purpose”? What could be improved?

The eight studies presented herein aim to cumulatively address the aforementioned questions. This body of research addresses an important gap in the academic and practitioner arenas by devising a conceptual framework of good regulation and applying it to examine the health technology regulation continuum of market authorisation and coverage and reimbursement. To date, there has been limited research on what constitutes good regulation in the context of health technologies and the effectiveness or usefulness of the different policies, processes, and practices that policy makers (and other relevant stakeholders) adopt and exercise to meet this end. Moreover, as previously evidenced, few studies have investigated the regulation of medical devices.

Key propositions underpinning the thesis

A number of key propositions underpin the thesis and provide a guiding foundation for the individual papers, including:

1. *Health technologies cannot be considered as one coherent group.* Research and discourse regarding health technology regulation tends to combine different types of technologies in one singular group or they only consider pharmaceuticals. Many technology assessment frameworks to aid coverage and reimbursement policy (and, to some extent, market authorisation) were developed with pharmaceuticals in mind, not other types of technologies, such as devices. Different types of technologies, however, possess unique attributes that can potentially impact their assessment, overall regulation, and use in and benefit to patient care.
2. *Market authorisation and coverage and reimbursement of health technologies should be considered as interconnected in order to effectively address some of the challenges of regulation.* Certainly those entities responsible for market authorisation and coverage and reimbursement have different missions and mandates – differences that arguably impact the similarities and differences between their respective evidence requirements and processes. Nonetheless, they share the central principle of balancing the benefits and risks in deciding whether a technology should be available for use in the health care system, and increasingly such decisions are being made based on similar evidence (even if assessments are conducted separately and with different endpoints). Figure 1 illustrates the potential synergies and overlap. Better understanding these synergies (and potential gaps or disconnects in meeting common objectives) is central to facilitating good health technology regulation.
3. *Important differences exist in the approaches to health regulation taken by different jurisdictions, but regulators share similar challenges in ensuring good regulation.* Countries assume different approaches to health policy, including health technology regulation, because of variations in their

organisation and financing, cultural traditions, and norms and values. Despite these differences, there are common challenges shared by regulators to ensuring effective health technology regulation, especially with regards to assessing the benefits and risks of technologies to inform decision making. Comparing regulatory policies provides a lens through which we can better understand how countries respond to such challenges (e.g. timely access to safe and effective therapies, managing limited budgets, needs for robust, relevant, and transparent information), which may highlight collective lessons or insights for best addressing existing and future issues that arise in regulating health technologies. According to Klein (Klein 1997: 1269), “the challenge to improving our capacity to learn from the experiences of other countries is to deepen our understanding of the respects in which they differ or are similar”.

4. *Evidence-based regulation is a dynamic, not static process.* Early proponents of technology assessment (and similar approaches) maintained that it would provide a more rational and linear process to decision making. Aside from whether such approaches do indeed meet these objectives, the process of technology assessment and applying evidence to policy is better characterised as dynamic, requiring regulators to modify and improve their policies and practices to reflect ever-changing circumstances. This is because technological development and innovation itself is complex and evolves rapidly, creating new products or incremental improvements to existing treatments on a frequent basis. Such changes result in new demands regarding evidence requirements, assessment methods, expertise to interpret and apply evidence, and patient access to care. Moreover, health technology regulation has become an increasingly visible, high profile and collaborative activity. By virtue of greater scrutiny as well as discussion and research on the topic, regulators are pressured to correct limitations and improve their processes. Also, the broader the range of stakeholders involved, the more policies and processes are required to evolve, in order to meet and address diverse and changing needs and expectations.

5. *The lines between state-centred and self-regulation with regards to health technology regulation have become increasingly blurred.* In any regulatory domain, including health technology regulation, there are generally state actors who are engaged in regulating, private sector actors who are being regulated, and third party parties who may provide input into how the particular area or sector ought to be regulated. However, the lines between these modes of regulation have become increasingly blurred in the health technology arena, with manufacturers and other stakeholders (e.g. patients) involved in the regulation process itself and interest and affected groups (e.g. medical associations, patient groups, other regulators) engaged in monitoring and enforcing good regulatory performance. In particular, such changes have resulted in “pre-emptive self-regulation”, whereby manufacturers react to concerns from regulators, other policy makers, and the public by directly engaging in regulation. The blurring of boundaries may lead to better regulatory outcomes.
6. *Attaining ‘good regulation’ cannot be considered in isolation of the particular organisation or agency.* Given the prominence of health care issues in society and the often very public nature of health technology approval and coverage decisions, broader political and social values and priorities influence the attainment of good regulation, despite the regulatory organisation’s or agency’s efforts. Such influences are increasingly persuasive given the increasing role of media, including social media, and the Internet in public life. Depending on the circumstances, these external forces can have a facilitative or inhibiting influence on good regulation.
7. *While often overlooked compared to effective regulatory output, “good process” plays an important role in achieving good regulation.* Although good process in and of itself is not sufficient to achieve good regulation, it is central in two main ways: 1) facilitates the likelihood of an effective regulatory outcome and 2) helps maintain the credibility and sustainability of regulatory decisions or policies when they come under question. This is particularly important in the case of health technology regulation, which, as previously noted, is increasingly complex, under external scrutiny, and

involves a broad range of affected stakeholders with often times competing interests. For example, if a regulator is criticised and challenged by a particular decision, the fact that it was derived by way of a transparent, inclusive, accountable, and independent process may protect against undue stakeholder influence and facilitate increased support and adoption.

The contribution of each study to the overall thesis and current evidence base

The thesis presents eight studies, five of which have been published (or are in press) in peer review journals (Sorenson 2012; Sorenson and Chalkidou 2012; Sorenson et al. 2012; Sorenson et al. 2013; Sorenson and Drummond 2014); two of the three other studies are currently under review. Taken together, the studies aim to explore the dimensions of (and issues raised by) the good regulation framework in addressing the aforementioned research questions (Table 2).

Table 2: *Overview and contribution of each paper*

Study	Title	Aims	Technology of Focus	Area of Health Technology Regulation	Main Research Question(s) Addressed	Dimensions of Good Regulation Explored
Study 1	‘Good regulation’ and public health protection in the European Union: Evaluating the European Medicines Agency	<ul style="list-style-type: none"> • Evaluate the regulatory performance of the EMA 	Drugs	Market Authorisation	Q1, Q2, and Q3	All (entire framework formally applied)
Study 2	Improving medical device regulation: The United States and Europe in perspective	<ul style="list-style-type: none"> • Provide a comparative analysis of medical device regulation • Explore key challenges facing device regulation • Discuss current and proposed reforms • Specify and explore actions to improve regulation 	Devices	Market Authorisation	Q1, Q2, and Q3	All
Study 3	Comparative analysis of pharmaceutical coverage and pricing in Europe: Policy levers and mechanism and insights for the United States	<ul style="list-style-type: none"> • Critically review pharmaceutical coverage and pricing policies across eight European countries • Explore evidence-based drug assessment processes in depth • Explore areas of regulatory improvement • Draw implications for US pharmaceutical coverage and pricing policy 	Drugs	Coverage, Reimbursement, and Pricing	Q1, Q2, and Q3	All
Study 4	Evolving reimbursement and pricing policies for devices in Europe and the United States and considerations of value	<ul style="list-style-type: none"> • Compare coverage, reimbursement, and pricing policies for medical devices in Europe and the US • Ascertain extent to which evidence of value is considered • Critically assess policy initiatives that have supported or could facilitate 	Devices	Coverage, Reimbursement, and Pricing	Q1, Q2, and Q3	Cost-efficiency and effectiveness

		value-based device access and payment policies <ul style="list-style-type: none"> • Identify measures for improvement 				
Study 5	The evolution and impact of health technology in Europe	<ul style="list-style-type: none"> • Investigate the evolution of HTA in Europe over last 10 years • Assess differences across countries • Evaluate impact of HTA on various outcome measures 	Both.	‘Coverage, Reimbursement, and Pricing	Q1, Q2, and Q3	All
Study 6	Decision making under certainty: An investigation of international coverage with evidence development policies in the context of medical devices	<ul style="list-style-type: none"> • Investigate adoption and use of CED policies in different countries • Explore application of CED to medical devices • Ascertain success of CED approach to date • Identify areas and strategies for improvement 	Devices	Coverage, Reimbursement, and Pricing	Q2 and Q3	All
Study 7	Valuing end of life care: The case of advanced cancer drugs in the United States	<ul style="list-style-type: none"> • Examine availability and use of advanced cancer drugs in the US • Review and analyse their value for money and other considerations of value • Assess key technical, political, and social challenges to assessing their value and applying evidence to decision making • Elucidate strategies to obtain greater value in advanced cancer care 	Drugs	Coverage, Reimbursement, and Pricing	Q2 and Q3	Due process, expertise and impartiality, cost-efficiency, effectiveness
Study 8	The politics of comparative effectiveness in the United States: Lessons from recent history	<ul style="list-style-type: none"> • Examine past federal attempts at technology assessment in the US • Identify reasons for success or failure • Distill political lessons for current and future CER efforts 	Both	Coverage, Reimbursement, and Pricing	Q2 and Q3	All

The first two studies provide an in-depth analysis of health technology market authorisation. The first study formally applies the entire framework in full to assess the EMA - the key regulatory body for authorising pharmaceuticals in Europe. In particular, it ascertains how well the agency functions to protect public health and meet other policy objectives and how it has evolved (and the measures used) to enhance its performance across the different dimensions of good regulation. This study represents the first in-depth evaluation, employing a conceptual framework of good regulation, of the EMA and its regulatory practices since its inception.

The second study critically reviews and evaluates existing medical device regulation in both Europe and the US. The analysis allows readers to understand the mandates, organisation, processes and policies of both systems (along with their respective strengths and weaknesses), compare and contrast proposed or adopted reforms, and identify further policy improvements to ensure that safe and effective devices are available to patients and that regulatory processes meet many of the procedural criteria (e.g. transparency, openness, impartiality) outlined in the framework. The comparison of Europe and the US is important in this context, as both jurisdictions have been heavily criticised for ineffective medical device regulatory practices and are undergoing reform (largely in parallel and with many of the same aims). To date, medical device regulation has escaped academic focus. As a result of recent reforms, there have been a couple of recent studies on US and European medical device regulation (Basu and Hassenplug 2012; Kramer et al. 2012) and focused commentaries on specific problems with existing regulation systems (Cohen and Billingley 2011; Freemantle 2011; Hines et al. 2010). However, this study is the first to provide an in-depth comparative analysis of the key issues involved in existing regulatory frameworks and how reforms might or could address outstanding weaknesses. In addition, it also offers practical actions to further strengthen medical device regulation in both jurisdictions.

The subsequent two studies, studies 3 and 4, focus on health technology coverage, reimbursement, and pricing. The third study provides a comparative analysis of the range of regulatory tools employed by European policy makers to meet key health system objectives regarding pharmaceutical coverage, reimbursement, and pricing. It delves into particular detail on the policies and

practices associated with assessing drugs (via technology assessment) to inform such determinations. In particular, the various dimensions of the framework are explored and how they facilitate or hamper technology assessment in particular and coverage and reimbursement more broadly. Because the US is a high spender on health care and has recently adopted health reforms aimed at improving the value of health care and “bending the cost curve” (CER being one such policy), the paper provides options for consideration on how the US might incorporate evidence from CER studies in existing coverage and reimbursement policies based on the European experience. This study contributes a comprehensive and updated comparative analysis of coverage, reimbursement, and pricing policies across Europe, providing an in-depth investigation of evidence-based approaches in particular. To date, most studies have focused on one particular regulatory tool (e.g. reference pricing) or country.

The fourth study focuses on medical device coverage and reimbursement (and, where applicable, pricing) in Europe and the US, with a particular focus on the use of evidence on the value of a technology to support such decisions (similar to the companion study on pharmaceuticals). It provides the reader with an understanding of institutions, policies, and processes involved in the coverage, reimbursement, and pricing of devices in both jurisdictions and the similarities and differences between them, in addition to their respective strengths and weaknesses for fostering patient access to cost-effective new devices. The paper also provides a critical analysis of various policy initiatives that have supported or could better foster value-based device coverage and reimbursement. This study is the first analysis on device coverage and reimbursement policies and from a comparative perspective. To that end, a key objective of this analysis is to provide an evidence base to stimulate debate on medical device coverage and reimbursement policy in the US and Europe, a topic that has received limited analysis or discussion. Compared to the third study on pharmaceuticals, this paper examines Europe as whole with specific country examples (rather than an in-depth focus on a particular selection of member states) to complement the other work (i.e. some of the countries examined in the third study use similar assessment processes for drugs and devices and the fifth study examines HTA processes and policies in Europe in more detail) and because robust data on national device policies are limited.

All together, the first through fourth studies provide an in-depth understanding of the institutions, policies, and processes involved in the market authorisation and coverage and reimbursement of health technologies, highlight the similarities and differences between the regulatory approaches taken for drugs and devices in Europe and the US, and elucidate shared (or divergent) challenges and opportunities to achieving good regulation. To that end, they also analyse the measures used or proposed in the two jurisdictions to better regulation, particularly in terms of the assessment and consideration of evidence on the value of technologies in coverage and reimbursement policy. In addition, the four studies aim to make a link regarding the important (and growing) synergies between market authorisation and coverage/reimbursement decisions in regulating health technologies.

The final set of studies (fifth through eighth) delves further into health technology coverage and reimbursement. In particular, the studies provide substantive analyses on the role and impact of evidence-based approaches to health technology coverage and reimbursement on achieving and maintaining good regulation. They also explore measures that regulators and other stakeholders have adopted (or might adopt) to improve policy in this area.

The fifth study expands some of the work provided by the third study by giving a comprehensive comparative analysis of how HTA systems and their role in coverage and reimbursement in Europe have evolved over the last decade. It identifies the key challenges and discusses policy actions taken by policy makers and others to better regulation in this arena. Part of the rationale behind these systems is that the approach will advance important good regulation goals, such as independence, transparency, impartiality and expertise, stakeholder representation, and more effective and efficient decision making. The study strives to provide a qualitative assessment, based on the available literature, of the usefulness and impact of technology assessment on these aims as well as on clinical practice, health and economic outcomes, and innovation. Earlier studies have examined the evolution of technology assessment over time in select countries (Banta and Oortwijn 2009; Drummond and Banta 2009; Eisenberg and Zarin 2002; Jonsson 2009; Luce and Cohen 2009; Perleth et al. 2009; Sigmund and Kristensen 2009; Weill and Banta

2009) or have looked at the development of HTA more generally. This study is the only existing work to provide a comprehensive retrospective comparative analysis of the establishment and use of HTA in several European countries. Indeed, it not only examines what changes have expired over time, but also identifies key trends and outstanding challenges associated with the overall approach.

The sixth study gives a more in-depth examination of coverage with evidence development (CED) policies discussed in the third and fourth studies. In particular, it explores the use of the approach across countries and namely with regards to devices. While the concept of CED and its use in a particular country has been discussed in the literature (Carino et al. 2006; Chalkidou et al. 2007; Chalkidou et al. 2008; Dhalla et al. 2007; Hutton et al. 2007; Levin et al. 2007; Levin et al. 2011; Longworth et al. 2012; Miller and Pearson 2008; Mohr and Tunis 2010; Tunis and Chalkidou 2007; Ramsey and Sullivan 2005; Trueman et al. 2010; Tunis and Pearson 2006; Walker et al. 2012), there are no existing empirical studies that comparatively examine CED across different countries. A few case studies on CED applied to a particular device have been published, but none that examine the use of CED for devices more broadly. Given that the evidence base for devices is often lacking at the time of coverage decisions (more so than for drugs), they are particularly viable candidates for this approach. This study addresses this gap, providing a comparative analysis of CED policies across a number of countries, including those that have not been covered to date in the literature (France, Germany, the Netherlands, and Switzerland), and their application to medical devices. The study is also the first to empirically gather and analyse stakeholder experiences with and perceptions of CED policies.

The seventh study provides a case example of the difficulties (technical, political, social) associated with assessing the value of new technologies and applying such evidence to coverage, reimbursement, and practice decisions. In particular, the analysis focuses on advanced cancer drugs, given their high expense, questionable economic value, and strong social and political support for patient access. These issues coincide with a particular need to consider the following dimensions of good regulation: due process, expertise and impartiality, cost-efficiency, and effectiveness. The existing literature on this topic has focused on two

principal streams of inquiry: 1) opinion pieces stating that drug prices are exorbitant and concerns that they provide questionable value for money and 2) limited studies examining oncologists' perceptions of the value of advanced cancer drugs (Fojo and Grady 2009; Meropol and Schulman 2009; Neumann et al. 2010). The paper provides a more comprehensive analysis, using an extensive body of multi-disciplinary research (health policy, health economics, political science, sociology), to critically examine the reasons why these therapies are notably expensive and the implications for affordability and patient access; how a variety of stakeholders (oncologist, patients, payers, general public) value advanced cancer drugs; and, the technical, political, and social challenges raised in ascertaining their value and accounting for such considerations in decision making. Overall, the study is intended to lend a better understanding of the current issues raised by technology assessment in this therapeutic area and how existing practice and policy might be enhanced.

The eighth and final study focuses on the politics of generating and using evidence in policy and practice. Resource allocation decisions are not simply technical in nature, but also political. The politicalisation of technology assessment, especially when used to inform coverage and reimbursement, is no more evident than in the US, where it has experienced a turbulent history, namely at the Federal level. Despite various attempts at institutionalising CER and similar approaches in US health care over the last several decades, research is lacking to understand previous attempts at adopting and implementing this type of research. There has been some research on select agencies or organisations, such as the OTA and the Agency for Health Care Policy and Research (AHCPR) (Bimber 1996; Gray, Gusmano, and Collins 2003), but no analyses have investigated the collection of entities and efforts over time. This study therefore provides an in-depth historical case analysis of the use of CER (and related approaches) in the US, focusing on efforts supported or adopted at the Federal level (successful and failed). The historical analysis elicits important lessons for the latest US investment in CER on new technologies and other health care services and programmes – PCORI. It highlights how good regulation or the lack thereof (across the six dimensions) can facilitate or hinder the usefulness of technology assessment in the US context.

Summary of the thesis methods

The principle methods employed in this thesis include qualitative documentary analysis and in-depth expert interviews. This section will first discuss the documentary research approach, followed by explanation of the usefulness of expert interviews. Subsequently, the methodological particulars to each study are outlined.

Documentary analysis

A broad definition of a document is a ‘written text’. A document is an important source of information, and such sources of data might be used in various ways in social research. Researchers (Bailey 1994; Denscombe 1998; Flick 2014) note that documents include institutional memoranda and reports, census publications, government announcements, proceedings, and policy documents, diaries, white papers, articles and papers, visual and pictorial sources and so on.

The documentary research method refers to the analysis of documents that contains information about a phenomenon of interest to study (Bailey 1994). The method is used in investigating and categorising sources, most commonly written documents, whether in the private or public domain (Payne and Payne 2004). This research method is just as robust and viable and sometimes more cost-effective than social surveys, in-depth interviews, or participant observation. As per Scott (1990: 34), a document “must be studied as socially situated products”. That is, documentary research is much more than recording or describing facts. It is a reflexive process in which the research confronts the “moral underpinnings of social inquiry” (Coles 1997: 6). “Documents do not stand alone” (Atkinson and Coffey 1997; 55), but need to be situated within a theoretical or conceptual frame of reference in order for its content to be understood.

Documentary research has been a staple of social research since its early inception. Along with surveys and ethnography, documentary research is one of the three major types of social research. The key issues surrounding the types of documents and the ability to use them as reliable sources of evidence must be considered by all who use documents in their research. Using this type of material in

a research study means that the documents are recorded as secondary data sources in the fact that they contain material “not specifically gathered for the research question at hand” (Stewart and Kamins1999: 11). Documentary studies also often call for originality in translating existing documents into trends or general concepts, and are particularly susceptible to alternative interpretations (Flick 2014).

In order to ensure proper handling of the documentary data collected for this thesis and guard against some of the limitations of the approach, Scott’s (1990) quality control criteria were employed. Such criteria include authenticity, credibility, representativeness, and meaning. Authenticity refers to whether the evidence is genuine and from reliable sources; credibility relates to whether the evidence is typical of its kind; representativeness refers to whether the documents consulted are representative of the totality of relevant documents; and, meaning refers to whether the evidence is clear and comprehensible. The criteria were applied flexibly and interdependently, as suggested by Scott, in that one criterion did not exclude another and the criteria were considered when selecting documents to review and include in each study.

Expert Interviews

Meuser and Nagel (2009) identify the expert interview as a specific form of applying semi-structured interviews. There are different perspectives on who is seen as an expert in the literature. As noted by Deeke (1995: 7-8), “...who and what are experts can be very different depending on the issue of study and the theoretical and analytical approach used by it. We can label those persons as experts who are particularly competent as authorities on a certain matter of fact”. Meuser and Nagel (2009) provide a more detailed characterisation, where an expert is a person who is responsible for the development, implementation or control of solutions/strategies/policies and/or who has privileged access to information about groups of persons or decision processes. They also outline different types of expert knowledge to be considered and sought in selecting and conducting the interviews. The three dimensions of knowledge include: 1) technical knowledge (specific knowledge of the field), 2) process knowledge (information on processes and procedures, typically for direct daily involvement in the field), and 3) explanatory

knowledge (subjective interpretation of relevance, importance, and potential influence of rules, ideas, policies, etc.).

Expert interviews can be used with different aims. Bogner and Menz (2009) suggest a threefold typology of expert interviews, where such interviews can be used for exploration, systematising, and theorising. Explorative interviews are helpful for orientation in a new field of study and to better structure hypotheses. The systematising expert interview focuses on the exclusivity or robustness of expert knowledge (i.e. person has expertise in a field or issue and likely operates in a position requiring such expertise) and is often used to obtain information that is not accessible otherwise. Here the focus is also on generating information for comparability and aggregation. Theory-generating expert interviews are appropriate when interviewees are considered more than an information source and the focus is on subjective aspects of an expert's knowledge, including motives and implicit beliefs about institutional or system functioning.

The aims and contents of this thesis align most fully with the systematising expert interview, as the objectives of the expert interviews used in particular chapters are to obtain specialised knowledge from deemed experts in the field – knowledge and information that is not readily available through the documentary analysis – for the purpose of aggregating and comparing the collected data to address the relevant topic(s) of inquiry and research question(s). However, it is common to include questions to capture more subjective aspects of an expert's knowledge in systematising interviews, which normally relate to the theorising interview. Because experts influence the establishment and adoption of regulatory decisions and internal and external stakeholder assessment of performance is important to good regulation, the interviews included some open-ended questions to gather experts' situated and subjective views on particular policies, processes, or dimensions of good regulation.

Table 3 outlines the advantages and disadvantages of the documentary research and expert interview methodological approaches.

Table 3: *Advantages and disadvantages of documentary research and expert interview approaches*

	Advantages	Disadvantages
<i>Documentary Analysis</i>	<ul style="list-style-type: none"> • Data readily available • Inexpensive and economical form of data • Researcher does not have to be present during data collection • ‘Non-reactivity’ – records unbiased by data collection process 	<ul style="list-style-type: none"> • Limited by the availability of the data • Inaccuracies in original material • Bias – ‘selective deposit or publication’ • Data studied may be out of context • Timely preparation before analysis
<i>Expert Interviews</i>	<ul style="list-style-type: none"> • Ability to obtain specialised knowledge from subject matter and/or process expert • Offer details and nuances not available through other research methods or data sources • Some differentiation of interviewees facilitates ability to obtain both high-level policy and more detailed procedural information 	<ul style="list-style-type: none"> • Knowledge obtained not neutral • Potential interaction effects (e.g. procedural not rigorously standardised; danger of anecdotal and illustrative ‘information’; not inter-subjectively repeatable)

Sources: Bogner, A., B. Littig, W. Menz (2009). *Interviewing Experts*. Basingstoke: Palgrave Macmillian; Dunn, W.N. (2004). *Public Policy Analysis: An Introduction*. Third Edition. Princeton, NJ: Prentice Hall; Flick, U. (2014). *An Introduction to Qualitative Research*. Fifth Edition. London: Sage.

To address the aforementioned weaknesses, Scott’s quality control criteria were applied to the data collection and analysis process, as noted previously. In addition, the various studies included a range of documents from a variety of sources to ensure accuracy, representativeness, and guard against selective deposit bias. In terms of the expert interviews, a range of interviewees were included to ensure the information obtained was as neutral or balanced as possible, and not merely anecdotal. In addition, the entire interview process was largely standardised, with formal and uniform procedures for inviting interviewees to participate, communicating the aims of the study(s), and posing questions (i.e. standardised interview guide).

Following data collection, the goal of the data analysis was to bring order, structure, and meaning to the mass of generated information. With qualitative data, structure must first be derived from the data, which requires systematically analysing it so as to tease out themes, patterns, and categories (Flick 2014). Punch (2005: 199) stresses, “there is no single right way to do qualitative data analysis – no single methodological framework”. Therefore, methods of data analysis need to be systematic and well structured. Miles, Huberman, and Saldana (2014) view data analysis as an interactive process comprised of three main components – data reduction, data display, and drawing and verifying conclusions. These processes transpire before data collection, during data collection as interim and early analyses are carried out, and after data collection when the papers are approached and completed.

Data reduction is the translation of information from one form to another to simplify storage, analysis, and dissemination to others (Miles, Huberman, and Saldana 2014). The main aim of this stage is to reduce the data without significant loss of information. Taking the gathered documents and reducing them to certain patterns and themes accomplished this objective. Flick (2014) refers to this process as “de-contextualisation” and “re-contextualisation”, which results in a higher level of analysis where deconstructing the data lends to the emergence of a larger, consolidated understanding of the issue, policy, or phenomenon under study. Data reduction was achieved in the thesis through editing, segmenting, and summarising the documents used to support each study and the overall thesis. Where appropriate, coding and memoing were subsequently employed to identify and note key findings, trends, and themes and, ultimately, to understand, conceptualise, and explain health technology regulation (particular policies and processes, evolution in regulatory priorities and tools, regulatory performance, and challenges and opportunities for attaining good regulation).

Data display is the process of presenting and analysing data, with narrative text being the most frequent form of display for qualitative data (Miles, Huberman, and Saldana 2014). Accordingly, the thesis primarily relied on narrative text, which enabled the documents and interview data to be organised and summarised in a

meaningful way. This was an iterative process, as the text consistently evolved during the writing process when new understandings or themes were developed.

The final stage, data drawing and verifying conclusions, actually occurs more or less concurrently with the earlier steps. Initial conclusions are noted throughout the research process, but are not finalised until all data is analysed and can be considered and contextualised as a whole. This final step of an analysis requires the researcher to interpret the reduced and displayed data.

Study 1

This study employed a comprehensive review of the literature on the EMA guided by the conceptual framework of good regulation. Relevant literature was identified through searches of bibliographic databases (PubMed, EconLit, Web of Science, Scopus) and the Internet (Google Scholar and Google for both published and grey literature [working papers, reports, agency committee reports, presentations, government and legislative documents]).

Study 2

The analysis presented in the paper is based on a comprehensive review of the literature. Relevant literature was identified through searches of bibliographic databases (PubMed, Web of Science, Scopus) and the Internet (Google Scholar and Google for both published and grey literature [working papers, reports, agency committee reports, presentations, government and legislative documents]). The paper also benefited from helpful comments from three anonymous journal referees.

Study 3

This study is based on a review and analysis of the available literature. Relevant literature was identified through searches of bibliographic databases (PubMed, Web of Science, Scopus) and the Internet (Google Scholar and Google for both published and grey literature [working papers, reports, agency committee reports, presentations, government and legislative documents]). The review focused

on Denmark, England, France, Italy, Germany, the Netherlands, Sweden, and Switzerland. These countries were selected because they represent a mix of different models of health care systems (Beveridge, Bismark, mixed models, centralised, decentralised) with divergent financing policies for pharmaceuticals. Moreover, the countries comprise the largest pharmaceutical markets in Europe. Therefore, the countries selected allowed for an in-depth comparative examination of European pharmaceutical coverage and pricing policies.

Parts of the paper also built on earlier work by the thesis author (Sorenson et al. 2008a), which examined the role of HTA in Europe, particularly with regards to coverage, reimbursement, and pricing policy. This work (a published book) was based on a comprehensive literature review and interviews of select experts in six EU member states (Finland, France, Germany, the Netherlands, Sweden, and the UK).

Study 4

The information presented in the paper is based on a comprehensive review of the literature. Relevant literature was identified through searches of bibliographic databases (PubMed, EconLit, Scopus) and the Internet (Google Scholar and Google for both published and grey literature [working papers, reports, agency committee reports, presentations, government and legislative documents]). The paper also benefited from the comments from two anonymous journal referees and helpful feedback from the journal's Editors. Moreover, the Commonwealth Fund provided valuable guidance on earlier drafts of the paper.

Study 5

The information presented in the paper is based on a review of the literature. Relevant literature was identified through searches of bibliographic databases (PubMed, EconLit, Scopus) and the Internet (Google Scholar and Google for both published and grey literature [working papers, reports, agency committee reports, presentations, government and legislative documents]). The review focused on England, France, Germany, and Sweden. These countries were selected because

HTA assumes a central role in decision making in each jurisdiction, albeit to varying degrees. HTA bodies in England and Sweden assume the most formal (regulatory) role, while they are largely advisory in France and Germany. Each country also employs somewhat different procedures and methods for assessing new technologies, and the role of different stakeholders in HTA varies across jurisdictions. The countries also represent different models of health care systems: England (Beveridge), Germany and France (Bismark), and Sweden (mixed model of elements of Beveridge and National Health Insurance systems). Taken together, the mix of countries allowed for an in-depth investigation of how HTA has evolved differently across different health care contexts.

In addition to the literature review, the paper also benefited from the comments from two anonymous journal referees and helpful feedback from the journal's guest Editors for that particular issue. Moreover, select academics/policy analysts in Germany and France provided helpful information and document translation for those particular countries.

The paper also built on earlier work by the thesis author (Sorenson et al. 2008a), which examined the role of HTA in Europe. This work (a published book) was based on a comprehensive literature review and interviews of select experts in six EU member states (Finland, France, Germany, the Netherlands, Sweden, and the UK).

Study 6

This study adopted a two-staged methodological approach. First, a literature review was conducted on international CED schemes. The review focused on CED policies in Europe and North America, namely Canada, France, Germany, the Netherlands, Switzerland, the UK, and the US. These countries represent a mix of health care financing systems and jurisdictions with more established and new CED schemes. In addition to these national CED schemes, the available literature on the CED approach more generally was searched. Relevant literature was identified through searches of bibliographic databases (PubMed, EconLit, Scopus) and the Internet (Google Scholar and Google for both published and grey literature [working

papers, reports, agency committee reports, presentations, government and legislative documents]). Second, semi-structured in-depth expert interviews (policy makers/HTA bodies, industry, and academics/policy analysts) were conducted to supplement the information gathered from the literature and to obtain information on expert experiences with and perspectives on the different national CED schemes and on the approach more generally.

Study 7

The information presented in the paper is based on a comprehensive review of the literature. Relevant literature was identified through searches of academic databases (PubMed, EconLit, Scopus) and the Internet (Google Scholar and Google for both published and grey literature [working papers, reports, agency committee reports, presentations, government and legislative documents]). The paper also benefited from discussions with US and UK health policy and economics experts associated with the LSE-Columbia Health Policy Group, feedback on drafts of the paper from experts in cancer policy and CER, namely Michael Drummond, Professor of Health Economics, University of York, and Kalipso Chalkidou, Director, NICE International and Visiting Faculty, Berman Institute of Bioethics, Johns Hopkins University, and the comments from an anonymous journal referee.

Study 8

The information presented in the paper is based on a comprehensive review of the literature. Relevant literature was identified through searches of bibliographic databases (PubMed, Scopus) and the Internet (Google Scholar and Google for both published and grey literature [working papers, reports, agency committee reports, presentations, government and legislative documents]). The paper also benefited from the comments from four anonymous journal referees and helpful feedback from the journal's Editor.

SECTION I: MARKET AUTHORISATION OF HEALTH TECHNOLOGIES

Study 1: ‘Good regulation’ and public health protection in the EU: Evaluating the European Medicines Agency

Introduction

With regulation perhaps the “area of greatest EU policy output” (Broscheid and Cohen 2007: 346), the European Commission has increasingly turned to supranational agencies to govern regulatory policy making. At present, there are 30 EU agencies operating in various policy areas. The growth of regulatory agencies is considered a significant development, forming part of the emergence of ‘distributed governance’ and an important element in a wider transformation of the EU’s administrative system (Chiti 2000; Egeberg 2006; Majone 1997; Trondal and Jeppesen 2008). In particular, agencies form a constitutive element within the so-called ‘new modes of governance’ approach of creating and enforcing rules at the EU level. This approach advocates a shift away from the traditional Community method of regulation¹³ to embrace softer, more responsive and reflexive modes, with the incremental and consensus-generating approach of the open method of coordination best conforming to this ideal (Trubek and Trubeck 2005).

Consequently, there has been considerable interest in and comparative research on the reasons underpinning the creation of agencies, their functioning, and implications for European governance (Barbieri and Ongaro 2008; Gehring and Kraphol 2006; Geradin and Petit 2004; Geradin et al 2005; Gilardi 2002; Gilardi 2005; Kraphol 2004; Rittberger and Wonka 2011; Thatcher 2011; Vos 2000). Despite a growing body of evidence in this area, there remains a paucity of evaluative research on the performance of individual agencies. To address this gap, this paper evaluates the EMA¹⁴, which is responsible for licensing new medicines in the EU, and has two principal aims. First, the paper strives to set out a conceptual framework for evaluating ‘good regulation’ and, second, to apply the framework to the EMA. Given the EMA’s primary function in assessing marketing applications for

¹³ Where the European Commission has the exclusive right of legislative initiative and decisions are taken by the Council in interaction with the Parliament, all under the supranational control of the Court.

¹⁴ The EMA, previously known as the European Agency for the Evaluation of Medicinal Products (EMEA), is also responsible for veterinary medicines, but we focus only on human medicines.

new medicines on the basis of efficacy and safety criteria, the analysis focuses on the agency's commitment to protect public health.

In meeting these aims, the paper adopts a somewhat different approach to evaluating regulation or regulatory agencies than pursued elsewhere in the literature. For instance, instead of developing and applying a set of detailed indicators or checklists and scorecards to measure regulatory performance (De Panizza and Visaggio 2006; Lee and Kirkpatrick 2006; Radaelli and De Francesco 2007; Renda 2006) or employing an econometric analysis to address specific dimensions of regulatory function (Gilardi 2003; Hanretty and Koop 2012; Montoya and Trillas 2009), the aims of this study is to apply a framework for assessing agencies around principles of so-called 'good regulation'¹⁵ and one geared less towards viewing market efficiency as the primary performance measure. Our approach is also less concerned with specific outcomes than with the EMA's overall work. Rather than focusing on a single aspect, such as approval guidelines, transparency or speed of decision making, as more descriptive work on the EMA has done to date¹⁶, the study evaluates the agency's operations in practice against its stated objectives.

The paper first considers the creation and regulatory objectives of the EMA. The subsequent section provides the conceptual foundation for the analysis, which involves extrapolating key themes from existing work on 'good regulation' and 'good governance' in developing an evaluative framework. The framework is then applied to the EMA and the paper closes with potential policy learnings and implications for the agency and pharmaceutical regulation more generally.

¹⁵ Wider debates on the different types of regulation are excluded, such as private versus public interest regulation, economic and social regulation, or self-regulation and de-regulation.

¹⁶ The EMA was subject to an audit/review by an independent consulting firm in 2000. Based on interviews with stakeholders, this evaluation was limited mainly to the agency's authorisation procedures and telematics. The report did, however, provide the basis for new regulation. (Internal Audit Service (2009) *Final Audit Report on Selected Administrative Procedures Supporting the Provision of Scientific Evaluation of Human Medicines in the European Medicines Agency*, April 30").

The EMA and public health protection

While the rich history underpinning the development and scope of EU public health competences¹⁷ cannot be reviewed in full here, it is worth highlighting several important milestones that contributed to the emergence of health protection as an area for agency authority and the creation of the EMA in particular. First, in the aftermath of the thalidomide tragedy, the European Community instituted the first legislation in the area of pharmaceuticals in 1965. Directive 65/65/EEC defined a medicinal product within the European market context and stipulated rules regarding the development and manufacture of medicines in the Community, along with initial guidelines for post-market monitoring. Importantly, it established safety and efficacy as the sole grounds for marketing approval, which still applies today. A second milestone was the 1975 establishment of a ‘mutual recognition’ procedure and the Committee for Proprietary Medicinal Products (CPMP)¹⁸, which aimed to speed up marketing applications for new medicines and to alleviate the burden of applications being made separately to each national authority. In particular, the committee was to act as the single authorisation and arbitration body for the community market. However, these procedures were not wholly successful in practice, as member states and the industry leaned toward continued use of the traditional national route for a variety of reasons (e.g. less cumbersome). Another attempt was made in 1986 through introduction of the Single European Act (SEA) and through the ‘concertation’ procedure¹⁹, in particular, to speed up the authorisation process. More broadly, the SEA effectively established the legal basis for the single market to take consumer health protection requirements into consideration, which was further supported by several subsequent Community developments (e.g. treaties and European Court of Justice rulings)²⁰. Additional legislation pertaining to good

¹⁷ See M. McKee, T. Hervey, and A. Gilmore, ‘Public health policies’, in E. Mossialos, G. Permanand, R. Baeten, and T. Hervey, *Health Systems Governance in Europe: The Role of EU Law and Policy* (Cambridge University Press 2010) and G. Permanand, *EU Pharmaceutical Regulation: The Politics of Policy-Making* (Manchester: Manchester University Press 2006) for a more detailed history of European medicines regulation.

¹⁸ Council Directive 75/318/EEC, OJ 1975 No. L147/1 and 75/319/EEC, OJ 1975 No. L147/13, respectively.

¹⁹ Council Directive 87/22/EEC, OJ 1987 No. L15/38.

²⁰ For further details on these developments, see, for instance, G. Permanand and E. Vos, ‘EU regulatory agencies and health protection’, in E. Mossialos, G. Permanand, R. Baeten, and T. Hervey, *Health Systems Governance in Europe: The Role of EU Law and Policy* (Cambridge University Press 2010); M. McKee and E. Mossialos, ‘Health policy and European law: closing the gaps’, *Public Health* (2006) 120: 16-20; T. Hervey, ‘EU law and national health: problem or opportunity?’, *Health Economics, Policy and Law* (2007) 2: 1-6.

manufacturing practice, labelling, patent protection, advertising and sales promotion, and wholesale distribution all followed within this free movement context²¹.

The aforementioned events led to legislation in 1993 that established the European Agency for the Evaluation of Medicinal Products (now the EMA). The EMA's creation was driven largely by the European Commission, along with informal networks of scientists; neither industry nor the member states were initially interested in a pan-European medicines licensing body. However, the then Directorate General for Industrial Affairs, now DG Enterprise and Industry, began championing the idea in the late 1980s. In particular, an agency was seen as a means to facilitate implementation of the SEA. Following several years of negotiation with European governments, the EMA opened in 1995 and subsumed the CPMP²². The Agency operates as a decentralised scientific agency and its principal responsibility is to evaluate all applications for marketing authorisation for new medicines²³ in the EU. It also monitors centrally-authorised products²⁴ and national referrals, develops technical guidance, and provides scientific advice to sponsors. With national medicines agencies (e.g. Medicines and Healthcare products Regulatory Agency, MHRA, in the UK) directly involved in EMA's processes, as further discussed below, the regulation of medicines remains a joint EU-member state competence. Such an arrangement reflects a 'hub and spoke' model of regulation (Groenleer 2009).

²¹ Commission Directive 91/356/EEC outlines the principles and guidelines of good manufacturing practice for medical products for human use, OJ 1991 No. L193/30; Council Directive 92/27/EEC on the labelling of medicinal products for human use and on package leaflets, OJ 1992 No. L113/8; Council Regulation 1768/92/EEC concerning the creating of a supplementary protection certificate for medicinal products, OJ 1992 No. L182/1; Council Directive 92/26/EEC concerning the classification of the supply of medicinal products for human use, OJ 1992 No. L113/5; Council Directive 92/25/EEC on the wholesale distribution of medical products for human use, OJ 1992 No. L113/1. Further, see L. Hancher, *Regulating for competition, government, law, and the pharmaceutical industry in the United Kingdom and France* (Oxford: Clarendon Press 1990); L. Hancher, 'The European pharmaceutical market: problems of partial harmonization', *European Law Review* 15 (1990): 9-33; and, L. Hancher, 'Creating the internal market for pharmaceutical medicines – an Echernach jumping procession?', *Common Market Law Review* 28 (1991): 821-853.

²² The CPMP, however, was transformed as a core scientific advising committee within the EMA.

²³ Includes medicinal products for human and veterinary use, including biologics and advanced therapies, and herbal medicinal products.

²⁴ Pharmacovigilance is part of the agency's mandate and, since 2005, it has maintained public access to the 'Eudravigilance' database, which is a network and management system for reporting and evaluating suspected adverse reactions during the development and post-approval phases of medicines. Accurate and timely communication of emerging data on risk is considered an essential part of the agency's pharmacovigilance program, with risk education, risk management, and risk minimisation activities being essential components. Such activities include use of risk management plans, Summaries of Product Characteristics (SmPCs), Patient Information Leaflets (PILs), patient alert cards, and periodic safety update reports (PSURs). In addition, the EMA also operates a Europe-wide clinical trials database to monitor adverse events and other relevant study outcomes.

Along with its establishment came a revamp of the earlier Community approval procedures, resulting in either a centralised or decentralised procedure to drug authorisation. The former represents the mandatory application route for certain products, namely biotechnology-derived products (including biosimilars), orphan drugs, and medicines for the treatment of HIV/AIDS, cancer, diabetes, neurodegenerative disorders, and, as of mid-2008, auto-immune and viral diseases. The centralised procedure is also open (voluntary) to products containing a new active substance not previously authorised in the EU and those that constitute a significant therapeutic, scientific or technical innovation or in cases where authorisation is in the interest of patients at the Community level²⁵ (EMA 2007b). Review of submitted applications is carried out by one of several committees²⁶, although the actual assessment work is undertaken by two national medicines agencies ('rapporteurs') working independently of one another, but under the oversight of the relevant committee. Following the committee opinion, the European Commission then issues a formal EU-wide decision, which is binding across member states. Alternatively, the decentralised procedure or 'mutual recognition' procedure, involves one member state granting a product a license, after which it can obtain authorisation in other countries without the need for separate national applications. This is the process for conventional products and allows member states to put forth a formal objection²⁷. Should a manufacturer seek to launch a product in only one member state, the application is simply made to the national agency concerned, and the EMA is not involved except to arbitrate in cases of dispute. Both the centralised and decentralised routes have a 210-day turnaround period from submission of application to an EMA opinion, which was instituted to facilitate the availability of new medicines in the EU.

Compared to other EU agencies, the EMA assumes a unique role in the Community. First, determining which medicines meet the standards of efficacy,

²⁵ To determine whether a product is innovative, the EMA considers if a new medicine a) provides a new treatment alternative to patients, b) is based on significant new scientific knowledge or on the application thereof, or c) was developed using a new technology or application of a technology. Even if a medicine does not constitute a significant innovation as defined, it may be of patient interest at the Community level if it addresses a particular health need, allows access to medicines, or provides another type of contribution to patient care in the Community.

²⁶ Committees include the Committee for Medicinal Products (CHMP), Committee for Orphan Medicinal Products (COMP), a Committee on Herbal Medicinal Products (HMPC) and, since 2007, a Paediatric Committee (PDCO).

²⁷ Member States may object and appeal on public health grounds, and the EMA has a protocol in place to consider such instances.

safety, and quality for EU market approval²⁸ requires a high degree of expertise and responsibility. This extends to the agency's post-market surveillance activities of already approved medicines, which requires the knowledge and proficiency to identify and evaluate adverse events and the authority to issue opinions on any changes to initial licensing agreements (e.g. recall or withdrawal of a medicine from the market). As these functions have immediate and long-term public health protection implications, these decisions carry considerable weight. This is heightened by the fact that the European Commission lacks the expertise and capacity to verify the agency's recommendations. The EMA's role, therefore, goes beyond simply influencing or informing the commission's decision making. It essentially instructs the commission on the decisions it should take; EMA opinions are frequently given to the commission as ready-to-deliver documents. Second, in keeping with the commission's interest in supporting the European pharmaceutical industry²⁹, the agency also serves as an agent of EU industrial policy. This has been a stated objective of the agency since 2004 (EMA 2004). Taken together, these various functions reflect a quasi-regulatory role for the EMA, and one that differentiates it from the more informational or guideline (advisory) roles of the other EU agencies and the regulatory powers of the FDA in the US³⁰. It is this role that we aim to evaluate the EMA, bearing in mind the overarching objective to protect public health. The following section establishes the parameters of our analysis and outlines our evaluative framework.

Toward a framework of 'good regulation'

If 'good regulation' is to be pursued, it is essential to measure both the quality and performance of regulatory tools, policies, and the institutions that wield them. However, this is a challenging exercise for a number of reasons. First and

²⁸These duties were laid down in 1965 in the first European Community legislation aimed solely at pharmaceuticals (Council Directive 65/65/EEC, OJ 1965 No. L22/369), and they remained the criteria against which all new medicines are assessed before being granted marketing authorisation.

²⁹There have been numerous initiatives to support industry, including the establishment of the EU Pharmaceutical Forum (<http://ec.europa.eu/pharmaforum>) and the Innovative Medicines Initiative (IMI).

³⁰The FDA carries similar oversight responsibilities as the EMA, although more expansive as the former also regulates medical devices, food stuffs, veterinary products, and cosmetics. However, the FDA holds greater executive regulatory power than the EMA, in large part due to the political and institutional constraints surrounding the comparative roles and interests of the European Commission and the member states in the EU polity. This relates to the imbalance or constitutional asymmetry between the Commission's economic and social policy competencies, and has been shown to have had an effect on the EMA's mandate and wider EU regime for pharmaceutical regulation. See G. Permanand and E. Mossialos (2005). Constitutional asymmetry and pharmaceutical policy making in the European Union. *Journal of European Public Policy*, 12: 687-709.

fundamentally, deriving at a consensus on the definition of ‘regulation’ is difficult. As Baldwin et al. (2011: pg. 2) note “there is no single agreed meaning of the term [regulation], but rather a variety of definitions in usage that are not reducible to some platonic essence or single concept”. This relates to a second challenge. Radaelli and De Francesco (2007: 83) highlight that “...the concept of regulatory quality is prismatic”, and they rightly question whether there is sufficient agreement on what constitutes good regulation and its use in different institutional settings. Indeed, conceptions of quality are likely to vary according to audience constituency, market position, or even discipline (Weatherill 2007). This is often due to divergent weightings of the various criteria underpinning good regulation. In the case of pharmaceuticals, for example, economists and policy makers may stress the pursuit of efficiency, citizens and politicians may emphasize the importance of accountability, transparency, and other process measures, and industry may place value on international competitiveness, predictability, or potential for market access or growth. Consequently, defining (and indeed measuring) good regulation in all cases is not possible. Nonetheless, it remains important that we seek to develop frameworks that can transcend multiple jurisdictions, sectors, institutional settings, and affected actors. To meet this aim, scholars and practitioners alike have produced a notable body of literature evaluating regulation at various levels, offering a diversity of concepts, ideas, and understandings. While a review of the literature is beyond the scope of this paper, major themes from several streams of research that are relevant to the analysis are considered and drawn upon to develop the framework.

Academic approaches³¹ to good regulation tend to be theoretical and less prescriptive in relation to the pursuit of specific outcomes than practical or empirical designs. The economics literature is focused largely on (economic) efficiency³² and high productivity (via encouraging investment and innovation) as the primary indicators of performance, especially when examining public service sectors (e.g. utilities, transportation) (Cubbin and Stern, 2004; den Hertog 2010; Fink et al. 2003; Peltzman 1989). Alternatively, the literature from political theorists and empirical-oriented political scientists emphasise the importance of achieving and furthering

³¹ The academic literature on good regulation covers a range of disciplines. Here, the focus is principally on the economic and political science literatures; other disciplines of inquiry include philosophy, socio-legal studies, and organisational studies.

³² Namely, that regulation is good if it is efficient in the sense that it maximises wealth.

accountability, transparency, legitimacy, and other procedural-oriented values (e.g. responsibility, control, openness, predictability, and responsiveness) (Black 2008; Johannsen et al. 2004; Majone 2001; Mulgan 2000; Scharpf 1999; Stern 1997; Powell and DiMaggio 1991).

Taking into consideration both economic and political science perspectives, Baldwin, Cave, and Lodge (2012) put forth the most comprehensive framework based on five ‘tests’ of good regulation and oriented toward “...those arguments that have a general currency when regulatory arrangements and performance are discussed in the public domain” (26). Such tests or criteria aim to transcend the biases of using efficiency (or any other one indicator) as a single measuring rod or justification for regulation. Moreover, the five tests are applicable to both the instruments of regulation and the regulatory authorities for executing them.

The first test, legislative mandate, judges regulators according to their success in achieving their mandates as authorised by Parliament. This presumes a public interest outcome to be served, and success requires that the regulator demonstrate achievement of its mandate(s), which allows for claims of public support. Further claims for support or legitimacy can be further made when the second test – formal accountability to democratic institutions – is achieved. The third test, due process, necessitates transparent and accessible processes. Here, the focus is on openness, fairness, and consistency of treatment, as well as the levels of participation regulators afford the public, consumers, and other affected parties. Due process effectively ensures proper democratic influence over regulation, thereby exercising a legitimising effect and securing public trust. Among other things, in practice this involves a strong ‘reason giving’ function - making decision and dissenting views available, delivering timely responses, granting access to documentation, and involving a wide range of stakeholders in various levels of regulation. Fourth, regulators are expected to have sufficient expertise to exercise judgement in a way acceptable to the public. This is particularly the case in situations where the decision maker has to consider a range of competing options or values and arrive at a balanced judgement on incomplete or shifting evidence. Finally, the regulator must be efficient. Claims to this test involve both productive efficiency and efficient regulatory outcomes. The former reflects whether a legislative mandate is

implemented at the least possible level of inputs or costs. The latter encapsulates whether the regulation under examination leads to results that are efficient³³, which unlike productive efficiency, is judged with a degree of independence from the mandate itself.

Majone (1996: 300) offers a similar view to Baldwin et al. (2012), where regulatory agencies require a combination of ‘control mechanisms’ to ensure their legitimacy, which he identifies as: “...clear and limited statutory objectives to provide unambiguous performance standards; reason-giving and transparency requirements to facilitate judicial review and public participation; due process provisions to ensure fairness among the inevitable winners and losers from regulatory decisions; and, professionalism to withstand external interference and reduce the risk of an arbitrary use of agency discretion”.

In parallel with academics, international bodies and intergovernmental organisations concerned with regulation at various levels, in particular the OECD and the World Bank, have published practical guidelines, typically employing scorecard approaches on regulatory reform, what constitutes good regulation, and what impact it may have in practice³⁴. Much of this work is concerned with economic regulation, often focusing on individual sectors or on issues associated with deregulation. As these guidelines are generally underpinned by an interest in promoting good regulation, via a better economic environment, we can identify several common features. Such elements include that any regulation 1) have a strong legal basis (regulators must be independent), 2) be clear and feasible to implement, 3) bring a net benefit, and 4) is efficient. These dimensions are echoed in various national level guidelines, with notable examples including the work of the Australian Office of Regulation Review (AORR 1998), the Council of Australian Governments (2007), the United Kingdom’s Better Regulation Task Force (BRTF 2000)³⁵, and the External Advisory Committee on Smart Regulation to the Canadian Government (CEACSM 2004). Again, these are largely economic focused, but in being issued by

³³ Outcome efficiency can be judged across two measures: allocative efficiency (whether it is possible to redistribute goods to increase the benefits to or welfare of any one consumer without making another consumer worse off) and dynamic efficiency (whether there is encouragement of desirable process and product innovation, and whether the system produces flexible responses to changes to demand).

³⁴ See Jaramillo 2013; Kaufmann et al. 1999, 2002, 2004, 2005, 2009; and, OECD 1997, 2005.

³⁵ See also the Regulatory Impact Unit’s work on the effective undertaking of regulatory impact assessments (RIU, 2003).

elected governments, they do take wider social concerns into account. This is also true of related guidance put forward at the European level. For example, the Mandelkern Report on Better Regulation (COM 2001a) served as a basis for drawing the Better Regulation Policy in the EU. In parallel, the 2001 White Paper (COM 2001b) on European governance³⁶ outlined five principles – openness, participation, accountability, effectiveness and coherence – aimed at engendering and maintaining trust in the way the EU governs. These principles are particularly important for those agencies whose remit carries direct social policy impacts.

The aforementioned publications were followed by several European Commission documents aimed at improving regulation in Europe, including proposals to simplify the EU regulatory environment (COM 2002; COM 2003) and several strategic reviews of the Better Regulation policy (COM 2006; COM 2008; COM 2009). In 2010, the commission shifted focus from “better regulation” to “smart regulation” (COM 2010a). Influenced by an emphasis on “smart, sustainable, and inclusive” growth set by the Europe 2020 strategy (COM 2010b), the Smart Regulation policy sought to place greater attention on the whole policy cycle – from design to implementation to enforcement to evaluation and revision. In particular, the policy focused on the importance of proper implementation and modification in light of experience to ensure that existing regulatory frameworks and policies are “fit for purpose”. It also outlined the need to enhance collaboration between EU institutions and member states and to make greater strides to open policy making to citizens and other stakeholders. Such objectives extended many of the aims of the Better Regulation agenda and aligned with the concepts of “responsive regulation” (Braithwaite 2006; Neilsen and Parker 2009; Ojo 2009), “risk-based regulation” (Baldwin and Black 2007), and “really responsive regulation” (Baldwin and Black 2007; Black 2008). Heavily influenced by the global economic crisis, the Smart Regulation policy has since been bolstered by the introduction of the commission’s Regulatory Fitness and Performance Programme in 2012. The programme aims to reduce regulatory burden and streamline regulatory administration by identifying

³⁶ Although distinct concepts, good regulation can be seen as an element or exercise of good governance, with similar criteria and principles applicable to both constructs. The OECD (2001) and Kaufmann et al. (2002) have both set out criteria of good governance, and then specifically linked them to regulatory agencies. Moreover, as much work on European integration considers regulation as the hallmark of the EU in terms of a *sui generis* form of governance (see Eberlein and Kerwer 2002; Majone 1996), the two concepts can be linked together in this context.

burdens, gaps, and ineffective measures, especially in terms of how EU legislation is implemented at the national and sub-national level (COM 2012a).

The framework principally draws upon Baldwin and colleagues' approach, and is bolstered with elements from the practitioner-oriented guidelines outlined above. While the former represents objective dimensions for assessing regulatory performance, it is largely indicative. As such, the latter substantiated criteria are therefore considered complementary to Baldwin et al.'s more overarching conceptual principles. Together, they help form a framework that encompasses both theoretical and empirical considerations for assessing regulatory performance. Table 4 outlines the integrated framework. In addition to the five tests indicated in Baldwin et al.'s framework, an additional criterion was added. In attempts to remain sensitive to the economic perspective given the EMA's two-part role, we endorse the differentiation made by the Canadian government between effectiveness and cost-efficiency. This allows a separation between regulatory efficiency and regulatory effectiveness, with the latter focused on the delivery of policy results as opposed to questions of cost and allocative efficiencies.

Table 4: Criteria of ‘good regulation’: An evaluative framework

Mandate	(External) Accountability	Due Process	Expertise & Impartiality	Effectiveness	Cost-Efficiency
Authorised legislative mandate to claim public support (1)(2)	Answerable to elected body representing the public (‘democratically responsive’) (1)(2)	Fairness, openness , transparency, inclusion of relevant stakeholders (1)(7)(8)(9)	‘Sufficient’ expertise in order to secure public support for exercise of discretion (1)(2)	Delivering intended results (1)(2)(4) or policy objectives, as based on standards and targets (6)(7)(8)	Costs and savings generated (1)(2) where benefits justify costs (4)
Able to balance and ensure independence and accountability (3)(4)				Promote innovation through incentives and goal-based approaches (4)	Minimise costs and market distortions (4)
Sound legal basis (4) and regulatory backing (authority) (5)	Able to justify decisions and be subject to public scrutiny (5)(7)(8)(9)	Ensure fairness amongst inevitable winners and losers (3)	Professionalism to withstand external interference and reduce arbitrary use of discretion (3)	Better than alternatives (6)	Understanding cumulative impact of policies (risk and problem awareness) and avoiding duplication and overlap (8)
Stated regulatory objectives (5) which are better than alternatives (6)	Subject to adjustment (6) ^a				

Proportionality (7)(9) and necessity (8)	Subject to appraisal by independent bodies (6) or other external scrutiny (6)(9) in order to avoid 'regulatory capture' (4)		Takes account of the cultures and understanding that operate within regulated bodies; responds to constraints and opportunities presented by institutional environments within which regulator acts; responsiveness to the logics of different regulatory tools and strategies; performance awareness and modification; adaption to change (10) ^b	Timelines (8)
Consistency with other (national and international) regulations/law/policies (5)(6)(7) and, in the EU, respecting subsidiarity(9)	'Reason-giving' and transparency to facilitate judicial review and participation (1)(2)(3)		Robust to errors (6)	
Based on verifiable performance criteria (4)(6)		Clear and practical for users (4)	Evidence-based decision making (8) ^c	
Enforceable (6)		Improving internal management and serving stakeholders (2)		

Sources: Authors' compilation based on: (1) Baldwin and Cave 1999; (2) Baldwin et al. 2012; (3) Majone 1996; (4) OECD 1997, 2003; (5) Kaufmann et al. 1999, 2002, 2004, 2005, 2009; (6) Australian Office of Regulation Review 1998; (7) UK Better Regulation Task Force 2000; (8) Canadian External Advisory Committee 2004; (9) Mandelkern Report, European Commission 2001a; (10) Baldwin and Black 2007.

a This is potentially limited by the mandate.

b Some of these issues also related to accountability.

c The Canadian Report considers evidence-based decision making as an element of the 'effectiveness' criterion, but we consider it applicable also to the requirements of 'expertise' and 'cost-efficiency'.

Before applying the framework to the EMA, it is important to acknowledge potential criticisms with this approach, mainly related to the fact that some of the criteria are over-lapping and in parallel may be difficult to measure. Indeed, Table 1 does suggest the potential for overlap in several areas, which are listed across multiple columns even where the initial source may only be applicable in one area. The potential for overlap, however, is not necessarily a weakness. Lodge (2004), for instance, argues that traditional discussions of accountability and transparency (as represented in the due process test) are inherent across the regulatory process and therefore should not be seen as separate dimensions of regulatory performance. A related concern may be that the individual criteria involve some level of trade-off, in particular between accountability and independence. Nevertheless, as any trade-off is context or case-specific and reflects the role played by the individual regulator, it becomes useful in assessing a given agency's position. Finally, there are challenges to measuring the various criteria. For instance, it can be difficult to precisely state and measure what fulfilling a mandate(s) entails, as most regulatory statutes grant regulators broad discretions and scope for exercising expertise and judgement in taking regulatory action. This can, in turn, also pose barriers to measuring effectiveness and efficiency. We deem the aforementioned points valid without being restrictive and, thus, consider them throughout our discussion. Therefore, we do not claim our framework to be definitive or without potential methodological caveats. However, it does encompass the main elements of good regulation relevant to evaluating a regulatory agency. The following section turns to this aim, where the outlined framework is applied to the EMA.

Evaluating the EMA

Legislative mandate

Table 3 demonstrates that a legitimate mandate and its various dimensions is perhaps the most widely-shared criterion of good regulation. While there are numerous facets to this test, they are all aimed at ensuring a clear, appropriate and (legislatively) authorised remit. Regulation 2309/93 committed the newly-established EMA to “protect public health and users of medicinal products” via an improved approval process for new medicines. To meet this objective, evaluations were

intended to meet specific criteria for marketing authorisation³⁷ and be more expedient and less contestable than the earlier mutual recognition and multi-state procedures.

Although the EMA's principal remit has not changed substantively since the founding legislation, the agency's mission statement has undergone several revisions over the years, expanding the EMA's objectives and becoming more specific over time. Moreover, starting in the early 2000s, greater emphasis was placed on the need to develop and ensure efficient and transparent review procedures, support the competitiveness of the European pharmaceutical industry, and employ pharmacovigilance activities to safeguard patient safety, which have been outlined in 'guiding principles' in support of the EMA's mission³⁸. In the last five years, additional focus has been placed on stakeholder engagement in the EMA's work and involvement in international regulatory standard setting and harmonisation (EMA 2010c). At present, the agency's statement, in respect to its main function, reads:

The mission of the European Medicines Agency is to foster scientific excellence in the evaluation and supervision of medicines, for the benefit of public and animal health.

Evaluation of the agency's mission and associated objectives (and their evolution over the years) raises a number of issues. First, it is not clear that working to "benefit public health" is the agency's main function in practice. Even on paper, the language is vague and inconclusive as to the centrality of this particular function; as currently written, it reads as more of a by-product of the agency's activities. To illustrate this point, the FDA's statement³⁹ is much more focused and, indeed, clearer:

³⁷Council Directive 65/65/EEC on the approximation of provisions laid down by law, regulation or administrative action relating to proprietary medicinal products, OJ 1965 No. L22/369.

³⁸In support of the mission statement, the EMA recently accompanied the mission statement with several key 'guiding principles', including: "we are strongly committed to public and animal health"; "we made independent recommendations based on scientific evidence, using state of the art knowledge and expertise in our field"; "we support research and innovation to stimulate the development of better medicines"; "we value the contribution of our partners and stakeholders to our work"; "we assure continual improvements of our processes and procedures, in accordance with recognised quality standards"; "we adhere to high standards of professional and personal integrity"; "we communicate in an open, transparent manner with all of our partners, stakeholders and colleagues"; and, "we promote the well-being, motivation and ongoing professional development of every member of the agency".

³⁹Statement abridged in respect of medicinal products.

The FDA is responsible for public health by assuring the safety, efficacy and security of human drugs... The FDA is also responsible for advancing the public health by helping to speed innovations that make medicines more effective, safer, and more affordable; helping the public get the accurate, science-based information they need to use medicines to improve their health; and, fostering development of medical products that respond to deliberate and naturally emerging public health threats.

Secondly, and more specifically, compared to the member states' own medicines agencies, the EMA's competencies appear more geared towards accelerating approvals as an end in itself (Lewis and Abraham 2001). To this end, the CHMP has a Scientific Advice Review Group, which provides applicants with scientific advice up to six years in advance of the submission of an application and to help in the presentation of supporting documentation towards achieving a positive opinion (so-called 'protocol assistance')⁴⁰. Such remit has historically gone beyond the support offered by the FDA, although the agency has strengthened its cooperation with industry in recent years, following urging from industry, patient groups, and other stakeholders for a faster and more predictable regulatory process. Currently, the FDA provides guidance to industry regarding what evidence (e.g. trial design, effects, etc.) they recommend to assure approval. This is particularly the case for priority drugs meeting unmet medical needs and "breakthrough" therapies.

There have been calls (mostly from industry and, to some extent, at the Commission level) for the EMA's cooperation with industry to be similarly strengthened, most notably in terms of streamlining regulatory processes and providing greater certainty for manufacturers regarding development issues in particular therapeutic and technology areas. While such calls may neglect the EMA's limited resources and capacity, they also seem to ignore a perhaps more fundamental issue as to whether this should in fact be the role of a medicines agency in the first instance.

⁴⁰As noted by Garattini and Bertele (2004), this is an irregular arrangement, as the CHMP decides on applications, acts as arbiter in disputes and where companies might appeal a decision, and provides advance help to the industry on pre-clinical drug development issues.

Nonetheless, this strong industry-supporting role was designed as an integral aspect of the agency from the outset. Public health interests were secondary to industrial policy and single market concerns throughout the policy process leading to Regulation 2309/93, where stakeholders representing the consumer or patient perspective were generally excluded from discussions, and suggested amendments from Community institutions, including members of the European Parliament and the Economic and Social Committee⁴¹, were often ignored by the European Commission in drafting its proposals (Permanand 2006). By comparison, the member states and industry were directly consulted and involved in developing the final model. It was also the internal market council of ministers rather than health ministers that discussed and approved the final legislation. Even before it commenced operations, the EMA was attacked by consumer and interest groups as operating as an industry-serving body (Orzack 1996). In particular, the agency was criticised for having a skewed mandate focused on accelerating market access rather than on ensuring more stringent approvals. This has also raised questions about the agency's ability to adequately balance and ensure independence and accountability.

The lack of clarity in the mandate makes it difficult to discern whether the EMA's stated principal objective is better than an alternative aim(s), which is another element in fulfilling the mandate criterion. To that end, the vagueness of the mission puts into question whether the agency is working towards important and verifiable performance criteria, especially those that are particular to and appropriate for public health protection. For example, although the various committees work to strict timelines for the completion of evaluations and decisions are required to be reached within 210 days, these are primarily speed of turnover targets aimed at facilitating the timely availability of new therapies. While such targets may indeed have an indirect impact on public health protection by ensuring patient access to needed beneficial new therapies, the primary impetus for accelerated review processes is more focused on the agency's role in cooperating with manufacturers and safeguarding innovation than that of public health protection.

⁴¹The Economic and Social Committee is the assembly of European 'social and economic partners' and is granted a reading of proposed legislation. Its role is consultative and its opinions and proposals are not binding.

Legislation in 2005 sought to improve on the EMA's initial 10 years of work (EMA 2005). However, the agency's mandate and the authority it has to exercise it were largely left unaltered. Insofar as the commission and industry equated quicker approvals and innovation-spurring intellectual property rights with public health protection, there were important new provisions on shortening approval times for more innovative therapies⁴² or in the case of public health emergencies, and for lengthening periods of data protection. There were limited public health provisions included of equal weight. Nevertheless, the new regulation did influence some important changes aimed at improving the agency's public health role, many of which were insisted upon by the European Parliament.⁴³ These included better packaging and leaflet labelling⁴⁴, allowance of conditional marketing authorisations, and increased funding for pharmacovigilance activities. Additionally, a 'compassionate use' clause was introduced, which enabled a provisional license to be granted via the centralised procedure for drugs treating chronic or debilitating disease and for which no viable treatment alternative exists.

In 2008, the European Commission published a major policy document on medicines, "Safe, Innovative and Accessible Medicines: A Renewed Vision for the Pharmaceutical Sector" (COM 2008). As intimated by the title, most of the proposals focused on facilitating a strong single market in Europe and enhancing the competitiveness of the pharmaceutical industry, rather than on public health. However, the document did emphasise the need to strengthen the EU framework on monitoring patient safety to decrease the number of medication errors and to provide patients with reliable and objective information on available medicines to aid decisions regarding their treatment.

As a continuation of the longer-term strategy introduced in 2005, the agency published a 'Road Map to 2015' document, which outlined its key strategic objectives over the next five years (EMA 2011a). In previous years, the EMA's responsibilities expanded as a result of access to the centralised procedures for both

⁴² Although, as noted, there is not a clear definition of what constitutes an innovative advance.

⁴³ It is worth noting that Regulation 726/2004 was approved via the Community's co-decision procedure, under which the Parliament and Council adopt proposals jointly, and where Parliament can, in extenuating circumstances, exercise a veto. Co-decision did not exist when Regulation 2309/93 was passed, and the Commission was therefore able to push through its proposals via the internal market council.

⁴⁴ Included the use of Braille and the International Non-Proprietary Name (INN), a unique, global designation used to identify a pharmaceutical substance or active pharmaceutical ingredient.

generic/biosimilar and non-prescription medicines, in addition to legislation in the fields of paediatric and advanced therapy medicinal products. Recent EU legislation on pharmacovigilance and falsified medicines has further increased the agency's coordinating role in the international pharmaceutical arena, especially with regards to patient safety. Some of these changes have increased the involvement of the EMA in public health protection. Accordingly, compared to the 2005 strategy, the 2015 road map document assumes a greater public health orientation, with two of the main strategic areas focused on addressing unmet public health need and optimising the rational and safe use of medicines. The first aim, in particular, focuses on stimulating medicines development in the areas of unmet medical needs and neglected and rare diseases (e.g. antibiotics) and more proactively preparing for and reacting to public health threats, namely global pandemics. Regarding the latter aim, the agency expressed a clear commitment to further strengthen post-market regulation by requesting sponsors to prospectively collect real-world data and strengthen synergies with the FDA's post-market surveillance Sentinel Initiative. In parallel, under a new legal basis granted by the pharmacovigilance legislation, the EMA can now require post-market studies.

One longstanding issue is the lack of requirement that new drugs demonstrate relative efficacy or benefit compared to similar products on the market. One consequence is that regulators and payers, while both aiming to avail medicines that contribute to public health, currently apply different approaches. Payers or HTA bodies, for example, require evidence on the comparative therapeutic benefit and costs of new medicines, where such information is used to support coverage, reimbursement and, sometimes, pricing decisions. Attaining this evidence, however, is made more difficult as a result of the type of evidence (i.e. efficacy of product alone) required by EMA and other licensing agencies. While the 2015 road map plan does not change the agency's position on requiring relative efficacy evidence for new medicines, it does acknowledge that it has a role to play in fostering closer interaction between both parties of the health care system, mainly through improving the information provided in the European Public Assessment Reports (EPARs) and engaging with HTA agencies in the early stages of drug development to provide joint scientific advice and debate evidence requirements.

Indeed, recent legislative developments and the growing complexity of new medicines have pushed the EMA toward greater focus on its public health function, at least in principle. This orientation is partly reflected in the recent addition of supporting ‘principals’ to its mission statement, which among them includes a statement that the agency is “strongly committed to public and animal health”. Whether or not the agency will effectively achieve these objectives is yet to be seen, especially in light of historical performance, where industrial objectives have often superseded the EMA’s public health protection responsibilities. Although DG Sanco, the European Commission unit responsible for health and social affairs, assumed responsibility for the EMA in 2009, the influence of meeting industrial policy objectives is well engrained within the agency’s orientation and operations. In order for the agency’s legislative mandate to be properly authorised in the manner that the framework implies, it must also have competencies to represent those that its mandate impacts. Given the EMA’s two-part role, it should therefore have relative capabilities in both industrial and public health policy. This raises questions regarding for whom the agency is and ought to be answerable to, and leads to discussion of the second evaluative criteria, the issue of accountability.

Accountability

As Table 3 indicates, accountability also features highly in conceptions of good regulation, including in the European Commission’s own guidelines. According to the White Paper on good governance, accountability is a ‘political principle’, which is important for establishing ‘democratic governance’. It further states that “....regulatory agencies must be accountable to institutions, operators concerned, and more generally the public” (COM 2001b: 10). Baldwin et al. (2012) argue that a regulator must be ‘democratically responsive’ to a body that is ‘properly representative’, in terms of the public trust. In other words, an agency must be answerable to an external authority in which the public has confidence. However, accountability is, in general, a contentious subject in the supra-national context. The unelected nature of the commission has led to a wide-spread notion of a ‘democratic deficit’ in the EU. Nevertheless, it is generally accepted that being accountable at the EU level means being accountable to the European Parliament, which comprises directly-elected representatives and exercises budgetary control. The parliament’s

representation on the EMA management board may help serve this accountability function, but as these representatives have limited direct contact with the parliament, it appears more cosmetic than substantive (Permanand and Vos 2010).

The public interest perspective, in particular, has posed a problem for the EMA. Various actors inside and outside of the agency have criticised the EMA for insufficient ‘democratic control’ and inappropriate accountability (de Andres Trelles et al. 2002; Garattini and Bertele 2001; Garatti and Bertele 2010; ISDB 2006). For instance, there was initial consternation about whether DG Enterprise was best placed to represent the health or public interest in medicines regulation. In the early 1990s, the axis of EMA accountability underwent considerable debate, most notably in 2002 with the parliament’s debate of the review of pharmaceutical legislation. At that time, members of the then CPMP wrote an open letter to parliament members requesting a transfer of responsibility of the EMA to DG Sanco. Their proposal was largely based on the notion that doing so would make the agency more accountable to public health interests and would also be closer in line with the model of accountability assumed by the FDA, which reports to the US Department of Health and Human Services (DHHS).

Although a transfer of accountability was not immediately pursued by the European Commission, the surrounding debates highlighted concerns regarding the lack of patient and consumer involvement in the EMA process, which was viewed as a potential deficit to achieving procedural accountability (and sufficient due process) and discussed further below. Consequently, following the introduction of the 2005 legislation, the EMA management board added two representatives each from consumer and medical associations. These representatives are generally appointed by the Council in consultation with the parliament from a list of candidates selected by the commission. The debates also underlined additional accountability issues with other stakeholders, namely industry and, in particular, around the fees paid by the industry to the EMA for review and evaluation of their products (i.e. ‘user fees’). While applicant fees are relatively common among certain national agencies, including the FDA (with a quarter of its total budget and 65% of its spending on reviewing drug applications coming from such fees), the EMA has generated criticisms regarding its financial dependence on industry. Currently, the EMA

receives 80% of its budget from user fees, with the remaining 20% coming from commission subsidies. The proportion of the agency's activities funded by industry is slated to increase, given that as of July 2012 the EMA started collecting additional fees from industry to support its growing role in post-marketing pharmacovigilance, including the maintenance of new committees (e.g. EMA scientific advisory committee, Pharmacovigilance Risk Assessment Committee (PRAC)) and strengthening the existing EudraVigilance database to handle a larger volume of post-market data.

Due process

The third test, due process, involves open, accessible, and fair regulatory processes, all of which are closely linked to transparency – an aspect of good governance that is closely related to independence and accountability. If a regulator is going to be successful in securing public trust and attaining credibility, it needs to be as open and forthcoming as possible with respect to its activities generally and of (scientific) decision making in particular. Among other factors, this means ensuring sufficient 'reason giving' by making decisions and dissenting views available, delivering timely responses, granting access to key documentation, and involving stakeholders in the regulatory process (Permanand and Vos 2010).

In the EU context, transparency most often means accessibility of documents, and, in this regard, the EMA is subject to the EU's legislation on public access to European institutions' documents⁴⁵ (European Parliament 2001). Its website, therefore, contains a considerable amount of information, covering both the science and the administration and operations of the agency. In particular, four main documents are released by EMA when a new drug is approved, including a 1) press release containing only general information; 2) summary of product characteristics (SmPCs), which is largely intended for prescribers; 3) patient information leaflet inserted in the drug package; and, 4) the EPAR summarising the documentation

⁴⁵ The basic principles on citizens' access to EU documents states: "Any citizen of the Union, and any natural or legal person residing or having its registered office in a Member State has a right of access of the institutions, subject to the principles, conditions and limits defined in this Regulation"...."Openness enables citizens to participate more closely in the decision-making process and guarantees that the administration enjoys greater legitimacy and is more effective and accountable to the citizen in a democratic system. Openness contributes to strengthening the principles of democracy and respect for fundamental rights as laid down in Article 6 of the EU Treaty and the Charter of Fundamental Rights of the European Union".

produced by the manufacturer and the procedures that led the CHMP to approve the new drug. In addition to positive opinions, the agency makes negative decisions public – a requirement instituted in 2005.

Despite the available information (and some movement to expand the range of accessible information), the agency falls short in certain areas. First, although a basic EU principle is to allow its citizens the widest possible access to the documents its agencies possess, as previously discussed, there are some exceptions. For instance, the EMA has and can refuse access to information if disclosure would threaten commercial interests, unless there is an over-riding public interest (EMA 2006a). Getzsche and Jorgensen (2011) attempted for three years to gain access to unpublished trial reports on the obesity drug, Orlistat, held by the EMA, before ultimately succeeding. The overarching reason given by the agency against release of the report was that it would undermine the protection of commercial interests. It also justified its decision by pointing to the resulting administrative burden of redacting the report(s) and that they would be worthless after removing any personal data. However, allowing researchers' access to unpublished trial reports is important for protecting public health, a point emphasised by Getzsche and Jorgensen in communications with the EMA, as such reports are notably detailed and provide more reliable data than published papers.

Available evidence also suggests that compared with unpublished trial protocols available at regulatory agencies, published papers often demonstrate widespread selective reporting of favourable results and underplay associated risks (Chan et al. 2004; Melander et al. 2003; Rising et al. 2008; Turner et al. 2008; Vedula et al. 2009). In addition, positive trial results are more often apt to be published than negative ones (Dolgin 2009; Rising et al. 2008; Stern and Simes 1997). In 2010, concerns regarding the benefits and safety of the flu drug, oseltamivir, instigated debate regarding the secrecy of the documents submitted for marketing authorisation of new medicines. The overall tenant of the debate was that a lack of data transparency makes it easier for companies to hide unfavourable data. Later that year, the EMA declared it would widen public access to documents, including trial reports and protocols (EMA 2010a). This aim was reinforced in the agency's Road Map to 2015 document (EMA 2011a) and in recent revisions made to the Clinical

Trials Directive by the European Parliament. In particular, the parliament's Committee on the Environment, Public Health and Food Safety called for full publication of all clinical trials data once marketing authorisation is obtained.

Second, the aforementioned documents published by the agency fail to provide comprehensive information that would be helpful for public health protection and for researchers to conduct additional analyses of trial data. For example, the SmPC does not mention when a drug is approved by majority vote, and does not give the reasons for the minority's opposition (i.e. attaining adequate 'reason-giving'). To provide meaningful assistance to prescribers, the objective of SmPCs should be to provide a more comprehensive listing of side effects, possible drug interactions in accordance with clinical importance, and documentation and standardisation of summaries for generic drugs.

The EPARs are also problematic and have been criticised as opaque, inconsistent, and even misleading (Barbui et al. 2011; Garattini and Bertele 2010; Gotzsche and Jorgensen 2011). In particular, Barbui et al. (2011) found that examined EPARs often lacked key data (or selectively included favourable published clinical trials) as well as consistent reporting of available evidence. These issues were particularly acute with regards to reporting of Phase III studies. EPARs also failed to capture the critical issues that the committee examined and discussed during their review and did not contain the initial reports submitted by the rapporteurs (i.e. two members of the committee who prepare a preliminary assessment report for the committee to discuss and approve) or the manufacturer's replies to any questions raised. This information would presumably play a central role in clarifying how the final decision was reached.

Third, the EMA cannot release any original documents that the manufacturer submits for the approval process. In contrast, in the US, the FDA can, under certain conditions, make at least substantial parts of the original documentation available to scientists, clinicians, or patients' representatives. Fourth, besides a comprehensive availability of information, issues of potential conflict of interest are present. In particular, except for press releases, all of the agency's documents are written in close collaboration with the manufacturer.

The above issues are largely attributable to the long-standing reporting structure to DG Enterprise. Industry considers it has the right to “commercial-in-confidence”, in order to protect the substantial investments made to develop a new drug. Any disclosure of data could benefit competitors and damage industrial interest and profits, which may subsequently reduce investments in research. This would also create a disadvantage for patients, who might in turn have access to fewer drugs (Garattini and Chalmers 2009).

Finally, as aforementioned, involving affected stakeholders in the regulatory process is central to good regulation (Baldwin et al. 2012). The EU has emphasised the need to involve a broad range of stakeholders, particularly civil society groups (e.g. patients, consumers, health professionals), as a central tenant of its good governance policy. To meet the European Commission’s aims, the EMA has a Patients’ and Consumers’ Working Party (since 2006), which provides recommendations to the agency and its human scientific committees on all matters of direct or indirect interest to patients in relation to medical products. Moreover, patient representatives are formal members of the agency’s management board and some of the scientific committees, and participate in medicines evaluation with the CHMP, among other agency activities. As previously discussed, industry representatives have a significant role in the EMA’s processes and in review of sponsored products, in particular.

However, while the noted participation from key stakeholder groups is a crucial component towards achieving due process and improving accountability, there are other related issues that must be duly addressed to effectively meet this aim. Firstly, representatives from stakeholder groups must be fully vetted to ensure no potential conflicts of interest exist. For example, during the first nomination process, DG Enterprise named the European Patients’ Forum. However, external reports pointed to the fact that the Forum not only receives funds from the pharmaceutical industry, but also benefits from the use of a public relations company that has several pharmaceutical companies as clients. The commission has since demonstrated preference toward involving the more prominent (and larger) patient groups, which has often been questioned, given that they are often financed, directly or indirectly,

by industry (Herxheimer 2003; Perehudoff and Alves 2011). In fact, a challenge with achieving due process (and, to some end, accountability) through expanded participation is that controversy will often attend to those individuals, groups, or bodies involved.

In addition to the issue of who participates lies consideration of the degree of representativeness of those involved. For example, the views and experiences of certain patient representatives may not reflect those of similar patients (e.g. those suffering from the same conditions(s)). This, of course, also extends to other actors, such as clinicians or scientific experts involved in the authorisation process. Demonstrating representativeness may include questions of competency, which is related to the next criterion, expertise and impartiality. Indeed, ensuring appropriate accountability encompasses meaningful participation of involved stakeholders, where proportional attention is given to all represented viewpoints and concerns. For instance, the well-organised manufacturer might manage to generate more effective pressure on the regulator than the heterogeneous group of consumers of these products (Abraham and Lewis, 2002: 78-82; Lewis and Abraham, 2001: 62-73).

Expertise and impartiality

An important contributor to whether a regulator, bureaucracy, or other arbiter exercises its duties in an effective manner is its impartiality and expertise (Thatcher 2002). This is particularly true for delegated agencies, where the efficiency of the regulatory process and the credibility of the agency depend on high quality data and cutting-edge expertise (Genoud 2003). Together, expertise and impartiality engender public trust and support, which allows for an agency to exercise discretion in their work (Baldwin et al. 2012). These elements also confer a level of professionalism to withstand external interference, avoid regulatory capture, and reduce arbitrary use of discretion in decision making (Majone 1996; OECD 1997, 2003). Securing a sufficient level of expertise also helps ensure that regulatory decisions are robust to errors.

The challenge for regulatory systems is how to construct a system of risk-benefit assessment that can accommodate the inevitably socio-political nature of the

required judgements. For example, the selection and interpretation of evidence crucially affects risk-benefit assessments, which are arguably influenced by the various social interests involved and the particular expertise of the arbiter. If some of the scientific experts on committees are themselves involved in clinical trials and the drug development process, they will likely identify more strongly with the aim of bringing new drugs to market. In such cases, it is difficult to discern when expertise or social interests leads to partiality, as different evidence could be selected and/or interpreted differently with similar levels of rationality and with divergent outcomes in terms of the risk-benefit assessment.

As previously noted, the EMA relies on several scientific committees to aid it in its reviews and decision making regarding approvals, with the primary committee being the CHMP. The members of the CHMP are largely nominated by the member states in consultation with the agency's management board, based on the strength of their qualifications and expertise with regards to drug evaluation. However, this has recently been modified, where, in choosing experts, the EMA considers individual expressions of interest from qualified experts. By doing so, the agency has opened opportunities for involvement to experts who may not be part of the "establishment" in their own country.

The committee's scientific advice role is unusual and expansive, as it decides on applications, appeals, participates in and coordinates importance pharmacovigilance activities, and provides advice to the industry on pre-clinical drug development issues. There are likely distinct advantages and disadvantages of this arrangement. For example, diverse involvement may deepen understanding and expertise of the broader drug review and approval process, but may also increase opportunities for conflicts of interest, given those involved in reviewing available evidence and making authorisation decisions duly input into providing pre-clinical development advice as well as appeals. However, the scientific assessment work of the CHMP is subject to an internal peer-review system to safeguard the accuracy and validity of opinions reached by the committee. Moreover, in the last couple of years, the EMA has strived to attain greater transparency about potential conflict of interest of its experts. In late 2011, the agency launched a database housing an electronic

declaration of conflict of interest for all its scientific committee members and other experts involved in the agency's work.

Much of the evaluation of medicines is carried out by the national medicines agencies on behalf of the EMA. While the EMA has required declaration of interests from its experts, national authorities have generally been much slacker in this regard. As a result of a critical indictment of the EMA from the European Parliament in 2011, the agency has been asked to disclose the terms of its agreements with the national authorities on such issues as the independence of committees, experts, and the evaluation process (Phillips 2011). At the time of the indictment, the parliament refused to sign off on the EMA's account, as a result of concerns that "there is no guarantee that the evaluation of human medicines is performed by independent experts" and that "some experts had conflicting interests". In particular, issues were raised about expert connections to industry⁴⁶. Additional changes since the indictment include a new screening process of the declarations of interest of the EMA's experts and committee members, including its management board, against their curriculum vitae and publication of the minutes of some scientific committees' meetings. As of February 2012, the agency also started requiring employees to file public declarations of interests and be assessed for conflict risk⁴⁷. The conflict of interest policy declares that pharmaceutical industry employment, a strategic advisory role, a consultancy, or financial interests as incompatible with expert work with the EMA, particularly with regards to the board.

Effectiveness

The effectiveness criterion can be best understood as whether the EMA has delivered intended results or policy objectives (Baldwin and Cave 1999; Baldwin et al. 2012; OECD 1997, 2003). In particular, this can be measured across two principal dimensions. First, the extent to which the EMA has contributed to the provision of the best possible scientific opinion for the centralised authorisation of medicines for

⁴⁶ Concerns were, in part, instigated by the activities of Thomas Lonngren, former Executive Director of the EMA, following departure from the agency. Lonngren's resignation came two months after he incorporated Pharma Executive Consulting, a consulting firm working directly with the private pharmaceutical industry.

⁴⁷ Employees are assigned to one of three conflict-risk classifications: Level one for no conflicts; Level two for minor, likely indirect conflicts, which may preclude the employee from full participation in some decisions; and, Level three for employees with direct conflicts of interest. The Executive Director may grant exceptions case-by-case to Level two or three employees.

the EU. Second, whether the EMA has achieved its mandate to protect public health by providing EU citizens with medicines fulfilling the basic requirements for quality, safety, and efficacy. Several of the practical frameworks for good regulation emphasise that achievement of these dimensions should be based on standards and targets (Australian Office of Regulation Review 1998; Canadian External Advisory Committee 2004; UK Better Regulation Task Force 2000) or ‘goal-based approaches’ (OECD 1997, 2003).

The standards that could be applied to address the first dimension, which attend to the quality and accuracy of the EMA’s scientific advice, involve the following: the number of reviews; input available, sought, and considered across experts and relevant stakeholders; and, organisation and responsiveness towards recent and future contextual challenges⁴⁸.

The widening of the scope of the centralised procedure has increased – in fact, more than doubled – the total number of initial applications for human medicines (average of 45 in the 2000-2005 period compared to 95 in the 2006-2012 period, with a peak in 2008) (EMA 2000, 2001, 2002, 2003, 2004, 2005, 2006c, 2007a, 2008, 2009, 2010c, 2011c, 2012a). The rise in applications is due principally to generic and biosimilar products, especially in recent years – a trend that will likely continue as the patent period starts to expire for an increasing number of drugs. The number of positive and negative opinions and withdrawn applications varies from year to year, and depends on a number of factors, such as the type and complexity of the products under evaluation, the robustness of the data in the application, and the type of applicant (EMA 2012). However, there has been a high average of positive opinions across time – about 77% of outcomes for new medicines were positive from 2004 to 2012 (EMA 2004, 2005, 2006c, 2007a, 2008, 2009, 2010c, 2011c, 2012a).

According to a recent survey conducted by Ernest and Young (2010), a large majority of the national authorities queried consider the output of the EMA centralised procedures to be of good quality. In particular, 87% of respondents deem

⁴⁸We assume Baldwin and Black’s (2007) definition of “really responsive regulation” to ascertain responsiveness, which includes the following aspects: accounts for different cultures, understandings, and attitudes that operate within regulated bodies and the regulated; responds to constraints and opportunities presented by institutional environments/external contexts; responsiveness to the logics of different regulatory tools and strategies; performance awareness and modification; adaption to change.

current timelines appropriate and less than one third think that some aspects of the process could be simplified. The majority (53%) are not in favour of an extension of the centralised procedures to other products. The experts also mainly rated EMA from good to outstanding. However, the respondents highlighted that the quality of the assessment may vary, depending on the national authority(s) and experts involved in the assessment team. Moreover, the lack of resources may impact the robustness or completeness of the assessment.

In a separate study, Downing and colleagues (2012) found that the EMA approved almost every application (96%) in a single review cycle, while only 62% and 69% of applications were approved by the FDA and Health Canada in a single review cycle, respectively. In both cases, more than 30% of applications required multiple reviews before approval. More than one cycle is typically required due to requests for additional statistical analysis, data collection, or sometimes new clinical trials. While a predominance of single cycle approvals may indicate greater efficiency at the EMA compared to its sister regulatory agencies, it could also indicate a tendency to emphasise speed of review over completeness.

However, single cycle reviews may be aided by the EMA's growing involvement in providing scientific advice and protocol assistance to sponsors during the research and development of new medicines. Scientific advice early on may contribute to the submission of better, more comprehensive and relevant data to the agency later on. As stated in the latest EMA annual report (EMA 2013: 32), "scientific advice is considered as a means to facilitate and improve earlier availability of medicinal products to patients and health care professionals....and, as a means to promote innovation and research". In 2012 alone, there were 339 requests for scientific advice and follow-up and 81 requests for protocol-assistance and follow-up (EMA 2013). The majority of the requests were received for products undergoing Phase III evaluation. Scientific advice appears particularly important for small to medium enterprises (SMEs) that may not have the in-house regulatory, financial and administrative expertise required to develop their medicine – 64% of registered SMEs requested scientific advice (EMA 2013). A study by Regnstrom et al. (2010) showed that seeking scientific advice from the agency and complying with it is associated with a greater chance of receiving a positive opinion.

Along those lines, experts contribute to different stages of the authorisation process, including in providing scientific advice (when requested by the company), in assessment teams, through the peer review process, in groups supporting the work of committees, and during discussions between committee members and member state representatives. The EMA has, as previously discussed, put a number of actions and requirements in place to safeguard conflict of interest and ensure a broad range of expertise is represented and involved. Nonetheless, existing evidence (Ernest and Young 2010) suggests that some specific technologies are less represented than others. For instance, less than one third of the respondents claimed to have some level of expertise in-house on gene or cell therapies or tissue engineering. However, national agencies working with the EMA often contract with external experts for clinical and scientific or research advice.

The final measure is the contribution of the EMA to its objectives in terms of effectiveness, which we argue can be seen as involving two different aspects: 1) whether the agency is organised (structure) in such a way that adequately meets current and future challenges and 2) whether the agency is responsive (procedure) to ongoing needs and challenges. The organisation of the EMA to effectively achieve its objectives depends heavily on its committees. According to the Ernest and Young (2010) study, the committee system is largely considered to be effective. However, the EMA has become more complex over time, through the addition of various committees, working parties, scientific advisory groups, and other ad-hoc groups. The number of committees and groups may indeed place the system under pressure, rendering coordination more difficult and potentially stymieing the efficiency and effectiveness of the EMA's activities. In recognition of this issue, in 2012 the EMA launched a new Scientific Coordination Board, composed of chairs of the agency's scientific committees, scientific advisory groups, working parties, and other relevant senior management staff, to ensure sufficient coordination between committees and that the standards they set for medicines development and evaluation are consistent. Nonetheless, the growing number and complexity of medicines are placing greater time and resource demands on committees. In addition, the rise in the number of committees and working groups increases the risk of duplication of efforts and wasted resources and potentially discordant standards or opinions. The new board

may be able to address such issues through regular review of committee activities and foster coordination and communication between these groups.

The development of a sophisticated organisational architecture and its associated activities can be seen as a reaction to the growing demands on the regulatory agency and the evolution of scientific advances and EU legislation. In this regard, the EMA – on the whole – has adapted well and reacted to the ever-changing regulatory and scientific landscape. For example, the agency has instituted a number of new initiatives and mechanisms to respond to the new EU pharmacovigilance legislation, as discussed further below. Moreover, as patients and the general public assume a more central role in their health care, the EMA has responded accordingly by involving these groups in the evaluation process and providing more transparent and accessible information about its activities and the risks and benefits of available medicines.

From a more scientific perspective, the EMA has produced guidance and other materials (e.g. reflection papers) to explore new scientific developments, such as biosimilars and advanced therapies, which not only stimulates stakeholder dialogue on these issues, but helps modify, where appropriate, existing regulatory practices to align with such advances and new challenges. Of course, some commentators would argue that the EMA has been slow or insufficient in reacting to existing and future needs. For instance, as previously discussed, EPARs could provide more comprehensive and helpful information than they do currently and it still remains challenging to obtain clinical trial data and evidence submitted by industry in support of market authorisation. Similarly, as more treatment alternatives are available to patients and providers and payers demand evidence of comparative effectiveness to support reimbursement decisions, the EMA could do more to support relative efficacy assessment of new medicines, as discussed in further detail below.

The second dimension of effectiveness focuses more on the public health protection aspect of the EMA's performance. Similar to the first dimension, there are a number of measures to ascertain the extent to which the agency is effective in protecting public health. These include the availability of high-quality, safe and effective medicines for EU citizens; support for development of medicines of major

therapeutic interest and need; impact of market surveillance and post-authorisation procedures; and, provision of quality information for EU patients and health care professional (aimed at health protection).

Although access to high quality, safe and effective medicines and their distribution do not fall strictly within the EMA scope of responsibility⁴⁹, the agency nonetheless contributes significantly to meeting this end through the centralised procedure itself, the quality of its assessments, guidelines production, and pharmacovigilance and other post-authorisation activities to monitor medicines use once on the European market. For example, the EMA produces a variety of guidelines, which contribute to harmonisation across member states (and at the global level) and access to medicines with a satisfactory level of quality, safety and efficacy. They also aid efficiency by making expectations more explicit for both applicants and assessors, which may reduce the evaluation workload. Such guidelines include scientific guidelines related to assessing the quality, safety and efficacy of new drugs, Good Manufacturing Practice (GMP) guidelines, Good Clinical Practice (GCP) guidelines, clinical trials guidelines, and pharmacovigilance guidelines, among others. In producing guidelines, the EMA has increasingly interacted with stakeholders to ensure their relevancy and usefulness through different mediums, including concept papers, focus groups, workshops, and consultation periods. Stakeholder interactions also include other regulators, namely the FDA; the EMA and FDA have launched collaborative GCP and GMP Initiatives, for example.

Nevertheless, there remain challenges. The ability to ensure the availability of the safest and most effective drugs for EU citizens may be hampered by EMA's current evidence requirements for new drugs. To date, the EMA only requires that new drugs demonstrate that they are efficacious and safe for a defined group of patients, but not compared to existing therapies⁵⁰. It, therefore, remains difficult for patients, clinicians, and other health care decision makers to determine whether a new drug is superior, equivalent, or inferior to existing treatment alternatives

⁴⁹ For instance, industry is not required to introduce a centrally approved medicinal product in all member states and distribution monitoring is under individual member states responsibility, apart from parallel imports monitoring.

⁵⁰ Only required when use of placebo is deemed unethical.

(Sorenson et al. 2011b; van Luijn et al. 2007), which may result in widespread use of potentially less efficacious and unsafe drugs, as highlighted by the recent case of the diabetes drug, rosiglitazone. The relative effect of rosiglitazone against pioglitazone emerged after years of widespread use (Juurlink 2010), where rosiglitazone was shown to increase the risk of myocardial infarction and cardiovascular death (Loke 2010; Nissan 2010). A lack of comparative efficacy data also makes it more challenging and time consuming for national HTA bodies and payers to ascertain the relative effectiveness of new drugs.

The fact that relative efficacy evidence is not required is arguably not efficient from a public health perspective, but it likely contributes to a quicker and less bureaucratic approval process. The recent road map plan fails to move towards relative efficacy requirements, although it does highlight the need to provide HTA bodies with transparent information to aid technology assessments and to engage with them from early medicine development through the medicine's lifecycle. Better information reporting in the EPARs, joint approaches to scientific advice, and mutual input on clinical guidelines are some of the key initiatives put forth to meet these aims (EMA 2011).

The promotion of the development of medicines of major interest has been an important aim of the EMA since its establishment. The creation of the Committee for Orphan Medicinal Products (operating since April 2001), the Paediatric Committee for Medicinal Products (operating since July 2007), and the Committee for Advanced Therapies (operating since January 2009) demonstrates the commitment of the agency to address important public health needs. Designation of Orphan Medicinal Product status, for example, provides applicants with various incentives to facilitate drug development and authorisation, including enhanced access to scientific advice and protocol assistance, fee reductions for many types of centralised activities, and potential eligibility for specific EU research funding. The year the Committee for Orphan Medicinal Products was established saw 83 application submissions; in 2012, submissions numbered 139. Interestingly, the committee has given very few negative opinions over the years, but this may be attributable to the relatively high application withdrawal rate, generally due to the medicine lacking the necessary criteria for the orphan designation. Most therapeutic areas have been covered by orphan product

designations, although the most represented areas are oncology and metabolic diseases (EMA 2012). Besides orphan, paediatric, and advanced therapies, in the last decade, the EMA has approved a number of important new medicines with public health benefits (Box 1).

Box 1: *Examples of important new medicines with public health benefits approved by the EMA*

- **Forxiga** (*dapagliflozin*): Treatment of type-2 diabetes mellitus - allows improvement of glycaemic control without increasing insulin secretion.
- **Constella** (*linacotide*): Treatment of moderate to severe irritable bowel syndrome (IBD) with constipation in adults - first medicine authorised specifically for IBD in the EU.
- **Vibativ** (*telavancin*): An antibacterial medicine treating adults with nosocomial pneumonia, known or suspected to be caused by methicillin-resistant *Staphylococcus aureus* (MRSA).
- **Dificlir** (*fidaxomicin*): A first-in class macrocyclic antibiotic intended to treat adults with *Clostridium difficile* infections, characterised by inflammation of the gut and severe diarrhoea.
- **Zytiga** (*abiraterone acetate*): An anti-cancer medicine with a novel mechanism of action, intended for use in combination with prednisone/prednisolone, for the treatment of metastatic castration-resistant prostate cancer in adult men whose disease has progressed on or after a docetaxel-based chemotherapy regimen.
- **Buccolam** (*midazolam*): The first medicinal product recommended for a paediatric-use marketing authorisation for the treatment of prolonged, acute, convulsive seizures in paediatric patients from the age of 3 to 18 months.
- Fourth and fifth influenza H1N1 pandemic vaccines intended for prophylaxis of influenza in an officially declared pandemic situation.
- A medicine for the treatment of moderate to severe manic episodes associated with bipolar I disorder, and another for the treatment of schizophrenia.
- **Revestive** (*teduglutide*): Treatment for adult patients with short bowel syndrome (SBS), a seriously debilitating condition – shown to additionally reduce parenteral nutrition requirements in patients with the condition.
- **Adcetris** (*brentuximab vedotin*): Treatment for Hodgkin's lymphoma and systemic anaplastic large cell lymphoma. Adcetris is an antibody-drug conjugate, which combines both an antibody and an active substance. The antibody can direct the medicine to a specific target on lymphoma cells, allowing a selective delivery of the active substance to tumor cells.

As previously noted, pharmacovigilance has received a high level of attention by the European Commission, other EMA stakeholders, and the media in recent years. Since 2001, the EMA utilises EudraVigilance to collect pharmacovigilance data on a daily basis from all member state authorities and from companies and clinical trial sponsors. EudraVigilance receives an average of 45,000 reports per

month and is analysed by EMA staff and relevant national authorities (on average 2,000 analyses are conducted per month) (Ernest and Young 2010). To further aid transparency and protect public health, in May 2012, the agency began publishing suspected side-effect reports for centrally authorised medicines on a public website. These reports come directly from EudraVigilance. The introduction of the database has coincided with year-on-year increase in the total number of adverse drug reaction (ADR) reports received by the EMA. For example, in 2012 the total number of ADR reports received increased by 34% compared with 2011, with a particularly notable increase in the number of reports coming from countries outside the European Economic Areas for centrally authorised products (60% increase) (EMA 2013). An increase of non-EEA ADR reports relates to the extended scope of ADR reporting as set out in the new pharmacovigilance legislation, particularly the expansion of reporting requirements from serious unexpected adverse reactions to reporting of all serious adverse drug reactions, and the inclusion of spontaneous reports submitted directly to patients and consumers without prior vetting by a health care professional (EMA 2013).

However, the EudraVigilance system is only as good as the reporting that supports it. While the increase in the number of ADR reports submitted annually indicates an enhanced commitment of stakeholders to provide data, the EMA has expressed concerns regarding the compliance of national authorities and the industry with reporting requirements and timelines (Ernest and Young 2010). One reason that perhaps hinders reporting is that EudraVigilance remains a fairly complex system to handle, dealing with substantial amounts of data, which may make it difficult to understand and manage. Alternatively, informants may be concerned about the potential consequences of their reporting for both themselves and their patients or consider pharmacovigilance processes too burdensome. The EMA has recently recognised some of these challenges and in 2012 produced the first set of guidelines of good pharmacovigilance practices (EMA 2012b), which contain a set of measures to facilitate the performance of pharmacovigilance in the EU. It also established a Pharmacovigilance Risk Assessment Committee that same year, which provides dedicated, expert oversight of all areas of EU pharmacovigilance. In addition, the agency created the Article 57 database – the first EU-level database of all authorised medicines. Once populated, it will serve as an important tool for regulators to

identify with greater accuracy and rapidity medicines for which public health concerns exist, whether due to pharmacovigilance or issues related to the manufacturing or supply chain of a medicine.

The communication of information to patients, health care providers, and the general public is a critical function to assist the EMA objectives in protecting public health. The EMA utilises a variety of mechanisms to relay information about new medicines to end-users. Product-specific information can be communicated through the product label or patient information package insert. Labels (often referred to as the summary of product characteristics, SmPC, in Europe) are sometimes updated to reflect new evidence (positive or negative) on a drug's safety and efficacy. Increasingly, regulators, including the EMA, have encouraged the inclusion of patient-reported outcomes (PROs) information in the product label or SmPC. PROs assess patient views on product efficacy, particularly related to symptoms, functioning, health status, quality of life, patient satisfaction, compliance, and treatment preferences – arguably all aspects contributing to public health. Such information is also advantageous to industry, as it serves to demonstrate a drug's 'added value' beyond safety and efficacy. A review of PRO labels for drugs approved in 2007 and 2008 showed that the EMA included signs and symptoms-based PROs in 55% of SmPCs authorised, activity limitations were included in 14% and health-related quality of life endpoints in 31% of the summaries (Caron et al. 2008).

Yet, SmPCs are not always of high quality or effective. An EMA study (2007b) found that mistakes are often made in the information provided in the SmPC, such as wrong shelf-life and batch number, wrong blue box (contains essential authorisation information). Such oversights may result in unsafe and inappropriate medicines use. The agency, however, has made strides in recent years to enhance the accuracy and usability of package summary information. For example, risk management plans are put in place (and increasingly common) in cases of insufficient patient information leaflets or inadequate labelling. In 2012, 128 and 190 risk-management plans were developed for initial-marketing authorisations and post-authorisations, respectively – this marks a 24% and 50% increase from 2011 (EMA 2013). The rise in risk-management plans highlights the impacts of the new

pharmacovigilance legislation. Similarly, through the legislation, the EMA has updated the summary product information template used by industry for the medicines they market in the EU. The updated template will make it easier to identify medicines that are subject to additional monitoring and to encourage ADR reporting. For example, as of September 2013, all medicines subject to additional monitoring will display a black inverted triangle in their product information summaries. In addition, the new template puts emphasis not only on the risks of taking the given medicine, but also on the benefits the patient can expect and provides concrete recommendations on the conditions of use of the medicine concerned.

The EPARs are also aimed to provide detailed information on newly approved medicines to a variety of stakeholders. However, as previously argued, the EPARs do not always provide sufficient information to inform patient and health professional decision making and help protect public health. EPARs are intended to be published immediately following market authorisation approval, but in practice only 28% are published within two weeks and 73% within a month (EMA 2011; EMA 2012).

Another avenue for patients, health care professionals, academic researchers, and the general public, among others, to obtain information is through formal requests to access EMA documents. As previously discussed, public access to documents has traditionally been limited. An analysis by Ernest and Young (2010) suggests that until 2010, only about two-thirds of the requests for access to documents were fully accepted. While there is no available evidence to indicate whether the agency's new 2010 policy on access to documents has increased the accessibility of documents, there has been a rise in the total number of requests for information. In fact, the number of requests almost doubled (108 to 207) between 2010 and 2011 (EMA 2013).

Finally, the EMA's website is an essential vehicle to provide stakeholders with key information about the agency, the regulatory process, and the products it evaluates. The agency has strived to simplify the website over time and make it more user-friendly. These efforts also coincide with the drive to increase the transparency

of the EMA's activities and, as a result, there are an increasing number of documents and other relevant materials available on the website.

Cost-efficiency

Regulation is cost-efficient when the output of regulation justifies the cost. In order to meet this end, it is essential to understand the cumulative impact of policies and to avoid duplication and overlap in regulatory activities. Moreover, achieving cost-efficiency must be based on meeting the other criteria previously discussed. Consequently, it is perhaps the most interesting (and complex) criterion from an evaluation point of view, in that it aims to capture the dichotomy underlying most regulatory policies – the tension between the public and private approach to regulation. Both approaches are premised on opposing interests between consumer (patients) and producer (industry), with the former focused on protecting societal, public health concerns and the latter maintaining that regulation is designed first to serve industry. It is this tension, in particular, that characterises many of the challenges raised in discussions regarding the four previous criteria. Consequently, there are definitional problems related to the cost-efficiency criterion, in terms of determining which objectives and whose needs are met first. In other words, how to balance and assess economic efficiency versus social objectives?

Historically, commentators have lauded the EMA for being more efficient – in economic terms – than the FDA, as evidenced by shorter time-to-market for new drugs and what some considered more 'streamlined' approval processes. However, this has changed in recent years, with recent studies highlighting the fact that it now takes longer for drugs, on average, to gain approval in Europe. For instance, Downing and colleagues (2012) found that for novel therapeutic agents approved between 2001 and 2010, the FDA reviewed applications involving novel therapeutics more quickly, on average, than did the EMA (or Health Canada, the Canadian drug regulator), and the vast majority of these therapeutic agents were first approved for use in the US. This trend also applied to the time of the first regulatory review. The median length of time for completion of the first review was 303 days for applications approved by the FDA, 366 for those approved by the EMA, and 352

days for those approved by Health Canada⁵¹. Similar findings were highlighted by Roberts et al. (2011), who found that between 2003 and 2010, the FDA approved 32 new anti-cancer drugs, while only 26 were approved by the EMA. The FDA not only approved more new cancer drugs than did the EMA, it approved these drugs more quickly. Of the 23 drugs approved by both agencies, the median time from marketing submission to approval was 182 days for the FDA versus 350 days for the EMA.

Certainly there are many factors that impact potential differences in timing of approvals between the various agencies. For instance, differences could be due to timing of entry into the different markets and the new information that becomes available as a result. Other factors may be differences in resources (funding, staff) to review new applications and the robustness of reviews. Therefore, it is difficult to ascertain whether the estimated times to approval for the EMA are a limitation or a success and the underlying contributing factors. Arguably the answer differs across stakeholder groups, where, for example, industry or patient groups would consider longer review times problematic. It may be the case that the agency is more closely scrutinising new drug applications or requiring more data to support approval, or it may be that the longer approval times are due to expanding responsibilities and activities, which would arguably slow review processes, especially if available resources were not increased to support a growing workload. Although this scenario may be somewhat more desirable from a public health protection perspective, it may put the EMA at disadvantage, in terms of efficiency and meeting important public health needs (by delaying approval of essential treatments).

Given that the agency has not made significant changes to its review processes or evidence requirements in recent years, any increase in time to approval may be due to changes to its scope and complexity of responsibilities (as well as the growing complexity of the products under evaluation⁵²). Indeed, the EMA's sphere of responsibilities has expanded over time, in line with new EU legislation. Most importantly, the centralised procedure now extends to orphan drugs, HIV/AIDS, cancer, diabetes, and mental health (neurodegenerative disorders) drugs, as well as to

⁵¹ However, if multiple cycles of review were required, the time of review was substantially longer for the FDA – a median time of 765 days.

⁵² Indeed, over time, particularly since 2010, there has been a decrease in the number of generic and hybrid applications and an increase in the number of applications for medicines with orphan designation. More complex applications under evaluation often require clarification and additional data prior to making a final opinion.

generics, biosimilars, and non-prescription medicines. Recent changes have transpired in the fields of paediatrics and advanced-therapy medicine products. As previously discussed, new legislation is under way, for example in the fields of falsified medicines and pharmacovigilance, which will further increase the agency's role in the pharmaceutical arena (EMA 2011a). Such changes have resulted in a marked increase in workload (EMA 2011a). A 2010 evaluation of the agency (Ernst & Young 2010) highlighted the fact that the main committees are overwhelmed with work and that consistency between the agency's numerous committees was a constant challenge. Consequently, the most recent Road Map document (EMA 2011a) emphasised the need to maintain efficiency in the agency's operations, placing it as "the primary focus for the agency over the next five years" (EMA 2011b). Attaining greater efficiency is seen particularly important, given economic pressures across the entire EU regulatory network. Many of the proposed strategies that may impact upon efficiency centre on greater collaboration with other EU authorities, national experts, and industry. For instance, the document outlines the intent to foster EU-wide pooling of expertise and data as well as close collaboration with the national competent authorities. The challenges to effective partnership are real, considering that more than 40 national agencies are involved; countries that differ not only by size of the country and associated resources, but also by their sophistication and experience in drug regulation.

In addition, one of the key strategic areas of the Road Map focuses on facilitating access to medicines, which addresses – among other things – time to market for new drugs. Suggested priorities such as promoting information and work sharing with other (global) drug regulators and 'staggered' marketing authorisation, in particular, could likely have a positive impact on cost-efficiency. Alignment between agencies would encourage a global approach to regulatory activities, such as the conduct of clinical trials, manufacture, and pharmacovigilance, which would not only bring greater efficiencies in EMA's operations, but also to worldwide pharmaceutical research and development more generally. Conditional or staggered authorisation would ensure that potentially beneficial drugs reach patients more quickly, while safeguarding public health by requiring additional evidence generation before a final approval decision is made. The 2015 Road Map defines a staggered approval approach for situations not covered by conditional market

authorisations (EMA 2011a). Approval would initially focus on restricted populations of good responders, but later modified as real-life data becomes available (EMA 2011a). The EMA appropriately acknowledges that this approach should not lead to reducing evidentiary requirement for first-time market authorisations, but rather allows more flexibility in addressing the particulars of a given drug and any uncertainties in the available evidence (Barker 2010). Clearly, the ‘staggered’ approach is still in its infancy, with considerable work to be done to develop and implement a viable framework.

Discussion

Given the proliferation of regulatory agencies across Europe, it has become ever more important to assess their performance in practice. However, to date, there has been a paucity of evaluative research on European agencies and in the health sector, in particular. This paper addresses this gap by evaluating the EMA, a highly influential agency within the health care arena. Rather than simply assuming a descriptive analysis of the agency’s performance, the study applies a framework grounded in academic and practitioner research on the key criteria reflecting or encompassing ‘good regulation’. Indeed, good regulation, and the role of the regulator, should be designed to enable ongoing appraisal of a regulator’s strengths or successes and weakness or failures. An external audit system of sorts is therefore needed to aid continuous reflection and improvement. If a regulator or regulation is to remain useful, it must be robust, flexible, and responsive. Moreover, periodic external assessment also helps guard the agency and its respective regulations from undue political influences or ‘regulatory capture’, or perceptions thereof (Dal Bo 2006).

As evidenced by our evaluation, the EMA has made strides, especially in recent years, to ensure or improve its attainment of good regulation across the various criteria. In particular, the agency has attained, overall, a more balanced approach to meeting both of the main tenants of its mandate – industrial support and public health protection. To be sure, industrial objectives remain central to the EMA; however, in recent years, there has been a shift toward greater attention and activities focused on its public health responsibilities. Such improvements include the

provision of more reliable and objective information on new medicines for patients (e.g. packaging and leaflet labelling), allowance of conditional marketing authorisations, increased regulation of and funding for post-market data collection and pharmacovigilance. In addition, the agency has increased its interactions and collaboration with other leading medicines agencies, such as the FDA, to better align regulatory processes and harness surveillance activities to improve patient safety. These changes may be attributable to the shift in oversight from DG Enterprise to DG Sanco, expanded responsibilities of the agency, and increased pressure from stakeholder groups to protect public health given the growing number and complexity of new medicines.

In addition, it can be argued that the agency has enhanced its accountability in recent years through greater representation of patients, consumers, and medicines on the management board and on other key committees involved in the approval process. This also contributes to meeting the due process criteria, through enhanced stakeholder representation and involvement and transparency of process. In parallel, the EMA has moved, at least in principle, to wider public access to documentation, such as trial reports and protocols, which also contributes to improved transparency. Some of these improvements will be facilitated by the revised Clinical Trials Directive, provided approval by the European Parliament in late 2013. Furthermore, the agency's cost-efficiency has also become of central importance over time, especially as it has expanded its responsibilities and activities. While actions to improve efficiency are under discussion rather than actualised, the EMA plans to meet this end principally through increased collaboration with national and international regulators and more flexibility in regulatory decision making for promising new innovations (e.g., 'staggered' authorisation). In addition, the EMA has largely been responsive to stakeholder demands, new legislation and expanded scope of responsibilities, and scientific advances.

Although the EMA has made certain strides towards ensuring good governance over the years, additional actions are needed to effectively meet this aim. One of the most significant areas requiring improvement is the lack of systematic provisions for obtaining important data to guide clinical practice and downstream research on the effectiveness and safety of new drugs. The recent revisions to the Clinical Trials

Directive are certainly important steps toward attaining greater transparency and helping independent, interested parties define the benefit-risk profile of new medicines before they are allowed on the market. The reliability and benefit of post-market studies will also be enhanced through access to original clinical data. Along with the changes proposed in the revised Clinical Trials Directive, the EMA should offer access to the rapporteurs' initial reports, the discussion between the CHMP and industry, and the minority opinions. In parallel, companies should be required to generate clinical study reports (describing the clinical trial and its results) using International Conference on Harmonisation (ICH) guidelines. This will help sponsors provide harmonised information and enhance transparency by greatly increasing the amount of data available to independent researchers and the public. Moreover, access should be prompt, ideally soon after the EMA's decision, and documents should be available in a user-friendly format. Other improvements in agency documentation could entail continuous review of guidelines to reduce any inconsistencies, ensure relevancy of information, and monitor their impact.

In line with the proposed changes to the Clinical Trials Directive, abolition of confidentiality would help make the system more transparent and enable clinicians and patient representatives to obtain information on new medicines and the associated approval process, establish greater public confidence in the EMA, and improve clinical research. Similar to the US, a distinction could be made between material that demands some degree of commercial secrecy, such as information on the production of the active ingredients and methods used for drug discovery, and findings from pre-clinical testing and clinical trials that are unlikely to be important for the competition.

Another key area of governance requiring improvement is protection against conflict of interest to better uphold impartiality of involved experts and, ultimately, more objective regulatory decisions and improved public trust. To this end, it may be advisable for the EMA to ban members of drug-industry-sponsored organisations from participating as patient and health care professional representatives on EMA's management board and/or scientific committees and serve as experts. There should also be a common protocol and criteria for appointing experts to ensure consistency within the agency and across member states. For example, since 2008, the FDA has

employed an 11-step algorithm for determining conflict and eligibility criteria for advisory committee participation. Moreover, the FDA recently adopted rules that conditions participation in advisory committee meetings by regular and special government employees on their acknowledgement that their financial interest information (range, not specific amount) and waivers will be made public. In conjunction with stronger participation rules, the EMA should instigate systematic and random checks to verify declarations of interest filed by experts. Beyond experts and employees, anti-conflict rules should also be extended to clinical investigators. The US Department of Health and Human Services, for examples, maintains a minimum disclosure threshold of \$5,000 for clinical investigators receiving Public Health Service funding and requires that any equity interest in private entities be disclosed. Of course, as the EMA looks to improve its own conflict of interest regulations, any new rules will only prove as effective as their enforcement.

As previously discussed, the EMA does not currently require evidence on relative efficacy for new drugs and, consequently, this information is often unavailable at the time of market authorisation. Van Luijin et al. (2007) estimated that comparative data was available for less than half of new drug approvals by the EMA and even in case where it was accessible, a limited proportion (~25%) were published and publicly available at the time of licensing. However, evidence on the comparative risks and benefits of new medicines is needed by a range of decision makers when a drug comes to market. Such information, for example, can help the EMA and other regulatory agencies to safeguard public health from inferior and unsafe treatments, ensure that HTA agencies and payers make funding decisions based on the best available evidence of different treatment options, and aid clinicians' and patients' understanding of what therapies work best and their appropriate position in the treatment pathway (Sorenson et al. 2011b).

In 2010, the agency did outline a role for relative efficacy evidence in cases where a new drug might be associated with safety concerns, and if treatment with a medicine of inferior efficacy might conceivably lead to significant, long-term or irreversible harm for the patient (EMA 2010b). However, relative efficacy evidence should not only be recommended or required in these circumstances, but for all conditions where alternative drug options (with similar mechanisms of action or

intended for the same indications) exist. Comparative information would demonstrate whether differences in efficacy or safety are clinically important, whether responses to a product by patients resistant to a different one are thoroughly assessed and documented, and reward those medicines that provide value to the health system (Naci et al. 2012; Sorenson et al. 2011b).

A number of steps could be taken to facilitate this aim. First, clear criteria (and any exceptions) need to be established. If evidence standards are set unrealistically high or criteria for relative efficacy study requirements are unclear, for example, manufacturers may prematurely terminate development programmes for potentially valuable drugs or authorizations may be delayed. Second, open discussion and agreement is needed between all relevant stakeholders on what type of study design(s) and associated standards can be accepted as fit-for-purpose for generating relative efficacy evidence (Cholski et al. 2010; Eichler et al. 2010). Third, increased investments need to be made to develop a robust clinical research infrastructure to support relative efficacy evidence generation, both in terms of establishing research priorities and funding studies. This aim might be supported by establishing an independent expert panel to recommend appropriate comparators, sample size requirements to demonstrate the margin of superiority, equivalence or non-inferiority between new and existing medicines, and options to improve the operational efficiency of both pre- and post-market studies (Eichler et al. 2010). Finally, the EMA could assist industry and researchers by developing guidance on how to best prospectively plan, conduct, and analyse such studies, as well as support a publicly-accessible database housing study protocol details and results (Eichler et al. 2010; O’Conner 2010).

The EMA’s pharmacovigilance activities have undergone notable changes and improvements in recent years. However, there is some indication that such processes, particularly ADR reporting and EudraVigilance, could be simplified to improve the rate of use and usability. In addition, the agency should utilise EudraVigilance to its full potential by developing better (and updated) quantitative and qualitative data analysis methodologies for capturing and interpreting potential adverse events.

Finally, it will be important for the EMA to evolve with and address challenges introduced by new therapeutic and scientific developments. For example, the issue of whether a patient population can be considered an orphan population may become more complex in the future. Indeed, the trend toward the development of targeted therapies and personalised medicine could lead to more and more segmentation of patient populations into sub-populations. The rationale for such segmentation should be carefully monitored, as these subgroups may end up meeting the criteria for orphan status, while being a sub-indication of a non-orphan disease. These types of situations will likely lead to an increase in EMA workload in the near future, which is problematic given the already increasing number of orphan applications and budget to review such medicines.

It is, of course, important to acknowledge any limitations to the analysis. First, this type of consolidated framework may not address all of the complexities of good regulation. It may, for instance, be considered too abstract to be practically relevant. Hood (2000) has demonstrated that even poor legislation can fulfil given theoretical criteria of good regulation and, therefore, suggests that they represent a wish list rather than tests. Similarly, Radaelli and De Francesco (2007) highlight the importance of context and the policy maker's own subjective interests in deciding on regulatory quality. However, it is precisely because of such points that the tested practitioner criteria were integrated within the six principal dimensions, and why the framework was applied to a specific case, the EMA. It can be argued that such criteria, when taken collectively, are useful. Adherence with only one or partial adherence with several would suggest that the regulation or agency under consideration is less deserving of support than one that adheres to several criteria, or adheres to them all to a greater extent. As La Spina (2003: 2) noted with regard to assessing regulatory quality, "...a decision will be legitimate, if the process that led to its adoption and its expected results are in line with such principles".

In addition, there may be concerns that the criteria employed are not universally applicable. As with other indicators (e.g. those outlined by Radaelli and de Francesco), the framework is indeed potentially useful in multiple contexts, not just the EMA or an equivalent regulator in the health arena. While we share Radaelli's (2000) broader concern that stakeholders uphold different criteria to

ascertain good regulation, it does not necessarily follow that this should preclude attempts at wider approaches. Indeed, as Table 3 demonstrates, the framework goes beyond a single perspective to offer a more encompassing social science-based approach, which draws upon inputs from practitioners and academics, and takes account of economic, legal and public policy perspectives⁵³.

Of course, the framework may benefit from further improvement. For example, while the indicators of target-setting and responsiveness are presumed to be captured within the various criteria used in the framework, namely effectiveness and, to some extent, accountability and cost-efficiency, it may be more advantageous to include them as separate criteria. Targets are themselves subjective constructs, however. Nonetheless, they can have an impact, particularly if developed externally and with input from a variety of stakeholders. Responsiveness, meanwhile, although presumably a part of the accountability and effectiveness criteria, is perhaps worthy of separate delineation in order to clarify what is required to meet this objective, especially given the recent emphasis of the European Commission on ‘responsive regulation’.

Conclusion

Although there is no perfect or all-encompassing framework of good regulation, it remains important that we pursue and develop mechanisms for assessing regulation in practice. Indeed, ongoing performance evaluation is a key part of the better or responsive regulation agenda (Bevan and Hood 2006; OECD 1999; OECD 2004). Furthermore, evaluative research supports better discourse on regulation, which in itself is a channel whereby regulatory reform gains legitimacy in both European and international circles. Radaelli and Schmidt (2004) posit that better regulation discourse enables policy makers to make sense of their reality – cognitive judgments about what is ‘good’ and what is ‘wrong’ in regulatory activities and governance. Because discourse is both coordinative and communicative (Schmidt

⁵³ This wider approach can be compared with the earlier-mentioned consultants’ review of the EMA, which either applied a strong economic (C/MA, 2000) or only examined the agency’s performance against one or two specific criteria, which mostly focus on outcome-oriented indicators (e.g. efficiency and effectiveness), such as the recent Ernest and Young (2010) evaluation.

2002), it may indeed begin with ideas and more normative activity of assessing, but ends in the more concrete arena of policy change and legitimacy.

In closing, this study has demonstrated that EMA's commitment to public health protection has historically been somewhat weaker or, perhaps, implicit with regards to its pursuit of its mandate and objectives, which also includes industrial policy goals. However, in recent years, the EMA has placed greater emphasis on meeting its public health remit, in terms of its pursuits and achievements, while duly attending to the aim to support research and innovation. Importantly, in parallel and often in interaction, the agency has continuously adopted new policies and processes to meet the other criteria of good regulation (e.g. transparency, stakeholder involvement, efficiency). As elucidated by herein, the dynamic nature of drug regulation and associated scientific advances will necessitate continuous evolution (and thus evaluation) of the EMA.

Study 2: Improving medical device regulation: The United States and Europe in perspective

Introduction

Medical devices are serving an increasingly central role in clinical practice, improving patients' health and quality of life. The medical device industry and the areas of patient care it touches have grown considerably in recent years. For example, the annual revenues of the US medical device industry rose from approximately \$85 billion in 2001 to \$146 billion in 2009 (Kruger and Kruger 2012). While part of this growth is due to the greater use of medical devices already on the market, new market entrants drove a large portion. During the 2000s, more than 30,000 medical devices were cleared by the US FDA's 510(k) pre-market notification pathway and more than 300 new devices received original pre-market authorisation (FDA 2011a). Along with the higher number of new devices, these technologies have become more complex.

The growing number and sophistication of medical devices have introduced regulatory challenges. Recent debates and events in the US and Europe have brought into question the effectiveness of the existing regulatory frameworks in both jurisdictions to ensure the performance, safety, and quality of new devices. In the US, for example, the Institute of Medicine (IOM) recently called for the FDA to eliminate its 510(k) clearance process, maintaining that it was an unreliable screen for the safety and effectiveness of devices (IOM 2011; Miller 2011).

Industry has generally taken a different stance, focusing on concerns that the US regulatory system is too slow, risk-adverse, and expensive. The European system is therefore often viewed as superior, given its somewhat faster regulatory process for devices and earlier access to some high-risk technologies (e.g. coronary stents, replacement joints) (Gottlieb 2011; Pollack 2011). However, European regulators have also faced criticism. In a commentary to the *British Medical Journal*, Freemantle (2011) asserted that the current European regulatory framework for medical devices, through the Conformité Européenne (CE) marketing process, is

inadequate to provide sufficient safeguards for technologies that affect morbidity, mortality, and health-related quality of life. The cited inadequacies include inferior regulatory evidence standards, non-transparent decision-making processes, and insufficient post-market surveillance to ensure devices' safety and long-term performance. The European Commission has echoed such concerns, stating a need to “adapt the European regulatory framework in order to secure patients' safety while favouring innovation” (European Commission 2011). Recent market recalls of articular surface replacement hip prostheses, PolyImplant Protheses (PIP) breast implants, and PleuraSeal for lung incisions, many of which were denied approval by the FDA, have further heightened concerns about current regulatory practices (FDA 2012; Heneghan 2011; Meier 2013).

Given that the US and Europe have recently introduced or are currently debating reforms of medical device regulation, it is an opportune time to examine the current regulatory policies and practices in both jurisdictions and identify areas for additional improvement. Despite the recent studies comparing medical device regulation in the US and Europe (Basu and Hassenplug 2012; Kramer et al. 2012), there is little in-depth analysis of the key issues in reforming the existing regulatory frameworks and strategies to be considered and employed to improve medical device regulation. The purpose of this paper is to fill this gap. First, the paper offers a brief comparative overview of medical device regulation in the US and Europe. Second, it examines the main challenges facing the regulation of devices, followed by an analysis of recent and ongoing reforms. The analysis closes with a discussion of additional policies and practices that could be considered in current reform plans, or in the future, to strengthen the regulation of medical devices in both jurisdictions.

Comparative overview of US and European medical device regulation

United States

The 1976 Medical Device Amendments gave the FDA primary authority to regulate medical devices and to substantiate “reasonable assurance of safety and effectiveness” before allowing manufacturers to market their products (GAO, 2006). This legislation has subsequently been updated, with the Medical Device User Fee

and Modernization Act (MDUFMA) of 2002, which established sponsor user fees for application reviews and set certain performance goals for the agency.

The FDA assigns devices to one of three regulatory classes based on their intended use, whether the device is invasive or implantable, and the risk posed by the device to the user. As Table 4 shows, the device class determines the level of evidence and evaluation required to demonstrate safety and effectiveness. Low-risk Class I devices are generally exempt from pre-market notification (510(k)) and FDA clearance before being marketed, although their manufacturers are subject to general controls, such as registering their name and products with the FDA. Medium-risk Class II devices usually are required to clear the 510(k) review process, which determines principally whether the new device is substantially equivalent to a legally marketed (predicate) device. Substantial equivalence means that the device performs in a manner similar to that of the predicate in its intended use, technological characteristics, and safety and effectiveness (FDA 2000). If a device is determined to be substantially equivalent, a clinical trial is usually not required to prove its safety or effectiveness. Other requirements (special controls) may be imposed, however, such as those for labelling requirements and post-market surveillance (Kramer et al. 2012). If the FDA deems a device to not be substantially equivalent, the manufacturer can petition for reclassification or file a *de novo* application.

High-risk Class III devices require closer scrutiny. These technologies are generally required to undergo the most formal review process for devices: pre-market authorisation (PMA), in which a device must demonstrate safety and effectiveness through the submission of clinical studies. Devices in this class that have been created from changes to previously PMA-approved devices may not be required to generate additional clinical evidence (Code of Federal Regulations 2012; FDA 2008). Novel devices without a predicate are automatically classified as Class III, regardless of their risk profile. But, if the device is classified as low to moderate risk, the manufacturer can apply for reclassification to Class II or I through the *de novo* process and need not undergo PMA, a process discussed further in subsequent sections.

To safeguard public health once a device is on the market, the FDA requires a range of post-market surveillance activities (Table 5), including adverse event reporting by manufacturers and user facilities (via the Medical Device Reporting [MDR] program) and post-market studies to ascertain and monitor the device's safety and effectiveness (Kramer et al. 2012). The agency also supports a number of surveillance data networks, such as MedWatch, the Medical Device Surveillance Network (MedSun), and the Medical Device Epidemiology Network Initiative (MDEpiNET), to identify and address safety problems and advance epidemiological methods for device surveillance.

Europe

Until the 1990s each member state had its own approach to regulating devices. To regulate a diverse and complex market and promote the “internal market” in Europe, new regulations, known as the New Approach Directives, were introduced by the European Council that defined the “Essential Requirements” to ensure devices' safety and performance (Kramer et al. 2012). These requirements apply to all countries. Therefore, if a device meets the requirements and receives a CE mark in one country, it can be marketed in all member states. A CE mark certifies that a device is safe and functions according to the intended purpose described by the manufacturer. Under these directives, devices are categorised into four classes according to the degree of risk associated with their intended use (Table 6).

Similar to those of the US, Europe's evidence requirements for market authorisation increase with the degree of risk associated with the device. Manufacturers of low-risk devices (Class I) are required only to self-declare conformity with the Essential Requirements to a national “Competent Authority”⁵⁴, such as the Medicines and Healthcare Products Regulatory Agency (MHRA) in the UK. More moderate- and high-risk devices (Classes IIa, IIb, and III) require a combination of clinical and non-clinical data on the device being evaluated. If available, data for an equivalent device already on the market may be submitted, if

⁵⁴ Competent Authorities are national bodies that designate independent Notified Bodies to monitor that device manufacturers conform to the Directives' requirements. Authorities also exercise oversight in that they monitor the work of the Notified Bodies. Within their remit, they can also monitor manufacturers' compliance with EU legislation as part of their market surveillance activities. To that end, Competent Authorities are also responsible for monitoring and reporting the safety of medical devices on the market.

available. Although clinical studies are generally requested for high-risk Class III devices, the evidence requirements are vague, not available to the public, and non-binding for manufacturers and studies need not be randomised (Fraser et al. 2011). For manufacturers claiming similarity to an existing product, a comparative literature review typically suffices.

Manufacturers of these devices select and pay one of about 80 largely for-profit, independent “Notified Bodies”⁵⁵ to evaluate their device and receive a CE mark. Award of a CE mark is based on an evaluation of safety and performance (that a device functions as intended), and not effectiveness (clinical benefit).

Once a device is on the market, manufacturers are required to report all serious adverse events to the Competent Authorities. In Europe, this information is collated into a central database, the European Databank on Medical Devices (Eudamed). In addition to vigilance information, Eudamed contains data on manufacturers, certificates issued, modified, suspended, withdrawn or refused, and clinical investigations. The use of Eudamed has been mandatory since 2011. Post-market studies also may be required if a device’s medium or long-term safety and performance are not known from previous use of the device or when other post-market surveillance activities would provide insufficient data to address risks.

⁵⁵ Notified Bodies ensure that device manufacturers conform to Directives’ requirements. In particular, they initially verify and evaluate manufacturers against EU legal requirements and standards before they market their products. Any changes to an approved design of a device must also receive further approval from the Notified Body. Not all bodies can verify technologies and not all Member States have Notified Bodies. National governments choose if they wish to designate a Notified Body or not. In doing so, they must ensure the body is capable of covering the products concerned and can monitor and evaluate the body and its work.

Table 5: US classification of medical devices and regulatory requirements for approval and post-marketing surveillance

Classification	Description	Pre-Market Requirements	General Time to Clearance/ Approval	Post-Market Requirements
Class I	These devices are typically simple in design and manufacture, and have a history of safe use. Device examples are tongue depressors, crutches, and scalpels. <i>No to negligible risk.</i>	Subject to the least regulatory control, most Class I devices are exempt from premarket notification and/or good manufacturing practices regulation, although some general controls apply (e.g. device registration and listing, labelling regulations).	Varies	Reporting of device safety and performance problems are mandatory for manufacturers, but voluntary for providers and users.
Class II	These devices are more complicated and are associated with a higher level of risk than Class I technologies and include endoscopes, infusion pumps, and condoms. <i>Low risk.</i>	Most Class II devices required to clear pre-market notification 510(k) requirements. In rare cases, clinical studies are required for a 510K submission. In addition, these devices maybe subject to other, special controls, including special labelling requirements and mandatory post market surveillance.	6-12 months	Use of MAUDE (Manufacturer and User Facility Device Experience database); MedSun (Medwatch adverse event reporting program); and, Medical Device Surveillance Network (network of facilities collecting data on device-related problems).
Class III	Devices belonging to this category usually support or sustain human life, are of substantial importance in preventing impairment of human health, or present a potential unreasonable risk of illness or injury to the patient. Such devices include coronary stents, defibrillators, and tissue grafts. <i>Medium and high risk.</i>	Most stringent requirements. Typically insufficient information exists to assure safety and effectiveness solely through general or special controls. Therefore, a premarket application (PMA) is required for Class III devices, which includes evidence from prospective, randomised control trials.	12+ months	Post-market studies are required for select devices, particularly Class II and III devices.

Source: Authors' compilation

Table 6: European classification of medical devices and regulatory requirements for approval and post-marketing surveillance

Classification	Description	Pre-Market Requirements	General Time to Clearance/ Approval	Post-Market Requirements
Class I	These devices are typically simple in design and manufacture, and have a history of safe use. They pose extremely little risk to the human body. Device examples here include reading glasses, thermometers, and examination gloves. <i>No to negligible risk.</i>	Manufacturer allowed to -declare conformity with the Essential Requirements.	Approval not required.	Manufactures are required to implement a post-market study and/or vigilance program according to national requirements, which includes reporting serious incidents to the relevant Competent Authority. Reports are synthesised in the Eudamed database.
Class IIa	These devices include short or long-term use of devices posing relatively low risk to the human body. Devices in this class include digestive catheters, infusion pumps, and powered wheelchairs. <i>Low risk.</i>	In general, manufacturers are required to submit a dossier of relevant, supporting literature (clinical and nonclinical) to substantiate safety and performance. Although there are general pan-European data standards, evidence requirements are fairly fluid, depending on what is submitted by the manufacturer and required or recommended by the relevant Notified Body.	1-3 months (+ any time required for the sponsor to address any deficiencies in the submission)	
Class IIb	These devices include those posing relatively high risk to the human body, including technologies such as respirators, dialyzers, and orthopaedic implants. <i>Medium risk.</i>			
Class III	Includes long-term surgically invasive devices that may endanger patients’ life. These devices include coronary stents. Also includes special Class III (AIMD) devices, which require a source of energy to function (e.g. pacemakers, defibrillators, cochlear implants). <i>High risk.</i>	Clinical studies generally recommended for high risk devices, but most are non-randomised and single arm (focused on demonstrating safety). Requirements are somewhat vague on the European-level and variable across Notified Bodies.		

Source: Authors' compilation

Comparing the US and Europe

The US and European approaches to medical device regulation have fundamental differences. For example, the FDA was established to promote and protect public health through the regulation of medical products, whereas the European system of Notified Bodies developed as part of a broader initiative to strengthen innovation and industrial policy across Europe. Notified Bodies therefore were not designed to function as public health agencies. Instead, the protection of public health lies largely with the Competent Authorities, with the extent of their role varying widely among member states. Kramer and colleagues (2012) argue that these differences help explain why the US and Europe have adopted different regulatory processes and evidence requirements for devices. For instance, in Europe devices must prove only that they work as intended, whereas in the US devices require evidence of effectiveness.

Another key difference relates to the organisation of the regulatory systems. In the US, the FDA oversees all regulation of devices. In contrast, the European system confers significant authority on a collection of governmental (Competent Authorities) and private (Notified Bodies) bodies to oversee device evaluation, market approval, and post-market surveillance. The US approach theoretically allows for better coordination and ease of enforcing regulatory requirements, although as mentioned earlier, some commentators believe that greater centralisation results in a rigid, lengthy, and costly regulatory process (Gottlieb 2011; Pollack 2011). While the more flexible European approach may grant faster market access to certain devices, it is not without problems. For example, evidence standards have been found to differ across Notified Bodies (Cohen and Billingsley 2011), which may encourage manufacturers to seek a CE mark from a less rigorous body. Decentralisation also hinders the collection and analysis of safety data, especially for rare, but life-threatening, adverse events, for which a substantial amount of patient information is required to detect potential problems (Thompson et al. 2011).

Outstanding challenges in US and European medical device regulation

Despite the differences between the US and European systems, both jurisdictions face similar outstanding challenges to effective medical device regulation. The next section discusses several issues needing improvement.

Establishing and upholding appropriate evidence requirements

Requiring sufficient evidence (and applying rigorous review mechanisms) to safeguard public health and certify effectiveness, especially for high-risk devices, is perhaps the greatest challenge currently facing both US and European device regulation. In the US, there are concerns that too many high-risk devices are evaluated through less rigorous review mechanisms (IOM 2011). Over the last 10 years, only about 2% of medical devices have undergone PMA (Sweet et al. 2011). A Government Accounting Office (GAO) study (GAO 2013) found that between 2003 and 2007, only 79% of Class III devices actually underwent PMA, with the remainder proceeding through the 510(k) pathway. Unlike PMA, direct evidence of safety and effectiveness is usually not required for 510(k) submissions, and only 10-15% of submissions contain any clinical data (GAO 2006). Furthermore, devices deemed substantially equivalent to devices previously cleared by the FDA do not need to go through the pre-market approval process, even if that previous model was never assessed for safety and effectiveness or recalled for a major safety defect (Ardaugh et al. 2013). One study investigating a cohort of high-risk recalls in the US showed that 71% of such devices had previously been cleared through the 510(k) process and another 7% had been exempt from review (Zuckerman et al. 2011). The greatest number of recalls was related to one type of device, automated external defibrillators, so these findings should be interpreted with caution. Besides the quantity of robust evidence are quality issues. Based on an internal analysis by the FDA, more than half of the 510(k) submissions it received had quality problems, including incompleteness or failure to address basic elements such as a description of the device and proposed indications for its use (FDA 2011b).

Even the PMA has challenges. FDA mandates only that PMA applications provide a reasonable assurance of safety and effectiveness (FDA 2012). The

evidence available suggests that this typically means applications were approved based on a single clinical study (Dhruva et al. 2009). In addition, only a minority of trials are randomised or blinded, use an active control group and hard end points, and are consistent in the way they account for patients and report data (Chen et al. 2011; Chen et al. 2012; Dhruva et al. 2009). Such standards differ for drugs, which are expected to show “substantial evidence [of safety and effectiveness]” and for which uncontrolled or partially controlled studies are not considered sufficient for approval (FDA 1998). An alternative perspective (often taken by the FDA, industry, and some analysts) is that devices are different from drugs and therefore should not be held to the same standards (Miller 2009). In particular, devices introduce challenges that render clinical trials less feasible. For example, for a surgical device, it is difficult to randomise patients for surgery or no surgery and/or blind patients or physicians. Moreover, many different types of devices make it difficult to apply one evidence standard to all devices.

Another issue arises from a stipulation of the Medical Device Amendments of 1976, which established varying safety standards for devices that the FDA deems as low, medium, and high risk, as previously discussed. The law applied immediately to new types of devices and directed the FDA to retroactively classify products that were already on the market when the law passed. This meant that Class I and II devices underwent review for substantial equivalence to devices already on the market. But, even though Class III devices were intended for PMA, they were allowed to receive review for substantial equivalence temporarily until the FDA down-classified them or required PMA.

Congress always intended Class III devices to undergo PMA, and in 1990 under the Safe Medical Devices Act, it directed the FDA to establish a timeline to complete the transition to PMAs for all devices that were to remain in Class III (IOM 2011). The FDA, however, still has not classified some of the “grandfathered” devices. As of early 2013, 19 different types of Class III devices are allowed to reach patients through 510(k) clearance (Ardaugh et al. 2013). Consequently, potentially high-risk devices continue to reach the market without ever being tested in humans. One such example is metal-on-metal hip implants. Ardaugh and colleagues (2013) traced the 510(k) history of the DePuy ASR XL Acetabular Cup System and found

that, in most cases, the predicates used for clearance were not metal-on-metal and were substantially different in design from the ASR XL or their clinical performance was poor. Almost a year after its approval, the ASR's high revision rate was discovered when it was compared with all other total conventional hip prostheses in the Australian Joint Registry. By this time, ASR's were implanted in millions of patients, many of whom suffered serious harm and, as a result, needed additional procedures to replace the device (Meier 2013).

In Europe, the majority of Class III devices need only to demonstrate their safety and performance, not that they directly benefit patients, and there are no requirements to verify the adequacy of submitted clinical data. In most cases, the submission of robust clinical data is limited, and often the evidence submitted is from laboratory testing, literature reviews, or small clinical trials (FDA 2012).

A less stringent pre-market review process increases the risk that later studies may demonstrate that the device has no benefits or identify important adverse events that did not emerge at the time of market authorisation. For example, although 10 times more drug-eluting stents are approved in Europe than in the US, many of those approved offer no advantage over other treatment alternatives for preventing restenosis or have worse outcomes than other stents (Di Mario et al. 2011). Other devices approved in Europe have been withdrawn from the market after later studies demonstrated poor performance or unexpected complications (Cohen and Billingsley 2011). The Notified Bodies' lack of uniform evidence requirements is another related concern. Such diversity results in regulatory unpredictability as well as inconsistent evidence standards being applied to similar devices.

Overall, US and European evidence requirements for devices introduce not only risks to patients, but also the wrong incentives to generate the needed evidence to better understand and evaluate the benefits and risks of new devices. Considering that manufacturers often take advantage of existing evidence from already marketed devices to gain approval for a new device, they are reluctant to undertake new clinical studies. In addition, because later devices may be able to claim equivalence, the first manufacturer to market does not have a very strong incentive to undertake extensive clinical studies. This may be exacerbated by the fact that many device

manufacturers are small-to-medium enterprises (SMEs) that often lack the requisite expertise and resources to conduct large clinical studies. Taken together, all these issues suggest that when a device (or procedure using a device) enters clinical practice, the information about its efficacy and short- and long-term safety is meagre. Accordingly, adoption of a new device may be driven more by marketing and the enthusiasm of clinicians than by evidence.

Monitoring and evaluating post-market device safety and effectiveness

Despite the variety mechanisms to collect post-market surveillance data in the US, such as the MedWatch and MedSun systems, the reporting of adverse events remains weak. Although by law, manufacturers must report deaths or serious adverse events, they are not required to if they decide that the events are unrelated to the device (Lenzer and Brownlee 2010). Furthermore, voluntary reporting by providers, patients, and health facilities is somewhat rare and may be subject to reporting biases. A 2009 government report (OIG 2009) pointed out that only 6% of adverse event reports come from health care providers and users. Several factors contribute to under-reporting, including the reports' voluntary nature, fear of litigation, difficulties in connecting problems with a device, and failure by patients and providers to understand their obligation to report (Malenka et al. 2005). Moreover, clinicians may not have sufficient time or support to collect and submit data routinely. The FDA's ability to detect potentially unsafe devices is further hampered by the fact that many post-approval studies required as a condition of approval are actually not conducted or are of such poor quality as to not produce meaningful post-market evidence (Lenzer and Brownlee 2010).

European post-market systems face similar challenges. Manufacturers are required to report adverse events to Competent Authorities, but the events' inclusion in Eudamed is dependent on the Competent Authorities, who are not mandated to report. Only a few national Competent Authorities provide the majority of adverse event reports and public notifications of device-related safety concerns (Kramer et al. 2012), and no mechanism is available for providers and patients to report adverse events. Eudamed allows information to be exchanged only between national Competent Authorities and the European Commission, and it is not available to the

public. In addition, Kramer and colleagues (2012) noted that the coordination and analysis of post-marketing reports from Eudamed are highly variable. Consequently, to date, Eudamed has had limited utility. While guidelines have been issued to address some of these issues, they are vague and remain at the discretion of manufacturers. Poor adverse event reporting, in addition to fraud and poor post-market regulatory oversight, was associated with the recent PIP implant recall (Heneghan 2011).

Without systematic post-market data collection, it is difficult for clinicians, other health professionals, and regulatory agencies to understand to whom health care is provided and the actual outcomes of particular procedures or the use of devices once they are on the market. This is particularly important in the case of medical devices, for which evidence regarding their performance is frequently limited at the time they are first used. Moreover, often only through the actual use of a device are unforeseen problems related to safety and performance identified and addressed (Cheng 2003). For example, an analysis of stent implantation between 2003 and 2004 using the Swedish Coronary Angiography and Angioplasty Registry (SCAAR) found that patients treated with drug-eluting stents (DES) had a higher rate of mortality than did those receiving bare-metal stents.³⁸ The findings caused upheaval, and prompted an immediate decline in the use of DES and an urgent review of their safety. A follow-up SCARR study (with data extended to 2010), however, found that the new generation of DES was associated with lower rates of restenosis, stent thrombosis, and mortality (Sarno et al. 2012). The difference in outcomes was largely explained by cardiologists' increased use of the device and the introduction of better stents.

Ensuring adequate and transparent information exchange on the benefits and risks of devices

The public's demand for accessible and transparent information about devices and the regulatory process has grown in recent years, and both the US and Europe have taken action to improve the exchange of information with stakeholders. For instance, the FDA produces publicly available information about its regulatory pathways for various device types and associated evidence requirements; publishes

advisory committee input on new devices; and, summarises its justification for its approval of high-risk devices and information about associated adverse events. The agency also requires the disclosure of any financial interests that a clinical investigator may have in a device or product sponsor. Although, the FDA does not publicly disclose this information, in its recent guidance on financial disclosure, the agency noted that it intends to provide information about the number of clinical investigators as well as financial information in the product reviews it posts for an approval decision (FDA 2013). In Europe, collected post-market data are shared with Competent Authorities, and individual Competent Authorities provide on their websites information regarding their operations.

Achieving an open and accessible information exchange still is elusive. In the US, much of a sponsor's application for a new device remains proprietary, as is information about applications not approved. Moreover, European Notified Bodies have no obligation to publish their decision-making process, the evidence provided by sponsors, or the basis for granting a CE mark. Additionally, post-market data are not shared with the public.

Current reforms to improve medical device regulation

The current regulatory systems for medical devices clearly must be improved. The paper next discusses several areas of reform that are under way or have been proposed.

Enhancing existing regulatory frameworks

The growing number and complexity of medical devices are challenging current regulatory frameworks. To address some of these challenges and those associated with the FDA's device review programs in general, in mid-2012 the US passed the Food and Drug Administration Safety and Innovation Act (Public Law 112-144) (Federal Register 2012). Among its various provisions, the law supports enhanced transparency and justification of significant agency decisions regarding device applications; a change in the agency's guidance when device modifications require pre-market notification before marketing; programs to improve the device

recall system; modifications of the *de novo* application process; new procedures to reclassify devices previously grandfathered into the system; and, mechanisms to enhance post-market surveillance, such as the inclusion of devices in the Sentinel surveillance system. While some of these actions are intended to make the regulatory process more efficient, such as changes to the *de novo* application process, others (e.g. device reclassification, Sentinel) strive to better safeguard public health.

Since the publication of the IOM report, the FDA has introduced additional measures to improve the existing 510(k) process, although it did not accept the IOM's overall recommendation to eliminate the program altogether. The FDA's initiatives include new guidances to improve the program's predictability and effectiveness (e.g. guidance to improve the quality and performance of clinical trials and clarify when changes in a device warrant a new 510(k) submission); additional programs to fortify the 510(k) systems, including analysing the use of multiple predicates; and, training for agency staff and industry on various facets of the program.

The European Parliament is currently considering proposals to reform the EU's legislation on medical devices and in vitro diagnostics put forward by the European Commission and the parliament's Rapporteur and Committee on the Environment, Public Health and Food Safety (European Commission 2012; European Parliament 2013). The commission's proposal offers insubstantial modifications of European device regulation. The parliament rapporteur and committee have called for far more oversight and transparency than the current system offers, with extra scrutiny of the highest-risk devices, including a more centralised pre-market authorisation process. Industry groups are fiercely debating the proposals, particularly the parliament rapporteur's and committee's measures, claiming that they would slow patients' access to beneficial technologies and hamper the "edge that industry has here in Europe" (Cohen 2013).

The latest parliament vote on the reforms sidestepped a centralised pre-marketing authorisation system, but supports a number of measures clarifying the roles and responsibilities of the involved parties (e.g. national authorities, clinical experts), fostering coordination and harmonisation across member states, enhancing

the oversight and standards associated with Notified Bodies, and increasing the transparency and traceability of devices. Among other things, this means that Notified Bodies will continue to grant market approval through CE certification, but will face increased oversight and quality assurance by the Competent Authorities and a new Medical Device Coordination Group (MDCG), especially for high-risk devices. The MDCG, composed of experts and representatives of relevant stakeholder groups, is intended to provide advice to the European Commission and to assist the commission and Competent Authorities in ensuring the harmonised implementation of the reforms. For instance, before a Notified Body can issue a certificate, the MDCG will have the ability to request a preliminary conformity assessment on which it can issue comments within a deadline of 60 days. A small group of independent scientific experts will support the MDCG in its decision making.

Although the Notified Bodies will retain much of their current authority, the new legislation does result in greater regulatory centralisation. The European Commission will be more involved in the review and approval of devices. For example, certain members of the commission, along with the MDCG and other experts, will advise on the designation of Notified Bodies and ensure that the member states charge comparable fees. The commission also will be responsible for maintaining Eudamed, which is central to the implementation of some of the new rules, particularly with regard to enhancing devices' transparency and traceability.

Other significant changes are requirements that certain devices (e.g. high-risk implantables) undergo assessment by specialised notified bodies designated by the EMA, the European regulator for pharmaceuticals. Notified Bodies will be expected to have permanent in-house competent personnel and technical and medical expertise related to devices and will be subject to assessment of compliance and ongoing monitoring. Manufacturers also will be subject to unannounced inspections and possible imprisonment if they commit fraud.

Strengthening pre-market evidence standards and requirements

The impact of the US reforms on device evidence standards and requirements is somewhat limited, with the most significant developments being changes to the *de novo* application process and the reclassification procedures.

In the past, the *de novo* process required manufacturers to submit a 510(k) application, which is exhaustively reviewed by the FDA before a device can receive a “not substantially equivalent” determination. If deemed not equivalent, the device will automatically receive a Class III designation. Only then can the manufacturer submit a *de novo* request to have the device reclassified from a Class III to a Class II or I designation. This complicated and somewhat cumbersome two-step process has resulted in the rare use of the *de novo* route and in unnecessary delays when it has been used. For example, only 54 *de novo* classifications have been made since the process went into effect in 1998, and once a *de novo* application is submitted, it takes the FDA an average of 240 days to review (Ladin and Imhoff 2010). Ladin and Imhoff (2010) found that *de novo* review times have increased in recent years and are sometimes longer than PMA reviews. The main reasons for the few *de novo* applications and the rise in review times are unclear. But, procedural inefficiencies are likely a cause, as well as the more complex devices being reviewed, the greater use of multiple predicates (IOM 2011), and the poor quality of applications (FDA 2011b). Consequently, some innovative, lower-risk technologies may have been inappropriately subjected to PMA approval or delayed market entry because of lengthy *de novo* review times.

The new *de novo* process outlined in recent reforms simply requires that manufacturers submit a request to the FDA for *de novo* classification, which will streamline the procedure by removing the requirement for 510(k) application and review. The FDA will have 120 days to issue a decision on classification. A recent analysis by the agency suggests that since 2011, the average 510(k) review time has decreased (FDA 2012), and the new process should raise the number of *de novo* applications and further shorten review times.

The reforms also make it easier for the FDA to reclassify “grandfathered” devices as either Class I or II or to call for a PMA application. The main change is that FDA will no longer be required to issue a reclassification regulation in order to reclassify a device, which used to require an economic review of the potential impact of reclassification. This process can take years to complete. As a result of the reform, the FDA can accomplish the same thing by administrative order, which should expedite the process. To date, six types of devices have been proposed for reclassification, including metal-on-metal hip implants, which are required to meet PMA review (Meier 2013). One area of uncertainty with the new process is that the reform called for all reclassifications to go before a panel. Consequently, it may now take longer for the FDA to down-classify certain devices and additional time may be needed to assemble the requested panels.

The European reforms generally uphold the safety and performance requirements outlined in the Essential Requirements under the current approach, even in the case of high-risk devices (European Commission 2012b). But, the reforms do require greater harmonisation of evidence standards across Europe (European Commission 2012b), and it is encouraging that the latest reform language suggests that the “clinical evaluation” of devices may include not only safety and performance, but also clinical benefits (European Parliament 2013).

Improving monitoring of post-market patient safety and quality of care

In the US and Europe, reforms have focused largely on improving post-market regulation to better safeguard patients’ safety and quality of care. Both jurisdictions introduced a unique device identifier (UDI) requirement to enhance the traceability of devices. In the US, device manufacturers will be required to place a UDI on the device’s label. Some devices will also need to be directly marked with the UDI itself. In addition, accompanying device information will be made available through the Global UDI Database (GUDID). As the FDA explained, the purpose of the UDI system is to provide speedy identification of devices associated with adverse events, assist with faster and more efficient resolution of device recalls, and deliver an easily-accessible source of definitive device identification.

The UDI system will allow devices to be incorporated into the Sentinel Initiative. Sentinel proactively monitors various data sources rather than relying on spontaneous reporting from manufacturers and health care providers, which will enable the more timely identification of safety issues. The system, however, was initially designed to track drugs (via a National Drug Code), and adapting Sentinel to monitor devices has been difficult because of problems with identifying specific devices in the available data. The UDI system will help address this issue by allowing information about specific devices to be integrated into electronic patient health records and health insurance claims, two of Sentinel's main data sources. UDI also will be able to improve other types of post-market surveillance, such as registries, and provide important information to and from relevant stakeholders as devices move from the manufacturer to the health system and eventually become part of patient care. With certain exceptions, implementation of these requirements will be based on device class (first applied to implantable, life-saving, or life-sustaining devices) over a period of five years from the Final Rule, which was recently released.

The goals and general requirements of the European UDI system are similar to those of the US to ensure a harmonised approach to device traceability and a globally accepted UDI system. The European approach will also have a Europe-wide UDI database. Most likely, the UDI information will be included in Eudamed (European Parliament 2013). It remains to be seen whether member states will decide to develop their own UDI systems, which could reduce the UDI's usefulness, but that seems unlikely if it becomes part of Eudamed.

In addition to the UDI system, the US reforms aim to improve the device recall system. In particular, the FDA is encouraged to proactively identify strategies for using recall information to improve the safety of devices and create tools to identify frequently recalled devices and the common root causes of safety problems. In addition, to ensure and speed up the completion of post-market studies, the FDA now requires manufacturers to submit study plans within 30 days of the agency's order and to initiate studies within 15 months.

Similarly in Europe, the reform proposals under discussion are considering several actions to achieve a more robust post-market surveillance system. The role of

Eudamed will be expanded. Member states will be required to submit information about the registration of manufacturers and devices; any CE certificates issued, modified, withdrawn or rejected; vigilance activities and outcomes; and, any clinical investigations. Manufacturers of high-risk devices will also have to submit a written report of the device's safety and performance and the outcome of the clinical evaluation, with the expectation that the summary be updated annually. The reform language also states that Eudamed should be robust and transparent and ensure access by the public and health care professionals to key parts of the database, such as vigilance and market surveillance information (European Parliament 2013). In addition, member states will use compatible harmonised reporting forms for adverse events and device deficiencies, and time lines for reporting will be established according to the severity of the event reported.

Additional directions for high-performing medical device regulation

While the current reforms in the US and Europe will go some way to address the current weaknesses in both systems, additional actions could be taken to further improve medical device regulation (Tables 7 and 8).

Pre-market evidence requirements

In Europe, there is no agreed-on requirement that the approval of medium- and high-risk devices be based on high-quality evidence of benefits that are relevant to patients. Patient safety could be improved by requiring an assessment of short- and long-term benefits and harms in well-designed RCTs, with the use of blinding and hard endpoints whenever possible. These requirements should be the same across member states (and Notified Bodies). There should be no region or body of least resistance, in which devices are approved more rapidly and on the basis of less evidence.

In line with more robust evidence requirements, European device reforms ideally would extend beyond enhanced oversight of the Notified Bodies. In particular, the reforms should contain a centralised review and approval process for high-risk (and, ideally, medium-risk) devices, with the approval of all other devices

going through the Notified Bodies as usual. It is encouraging that the latest reform proposals are moving in this direction by requiring specialized Notified Bodies to review certain high-risk devices. Moreover, it will be important to ensure that all Notified Bodies have enough in-house expertise to review an increasingly diverse range of devices. The new requirements for standardised processes, evidence requirements, and fees for Notified Bodies should protect against manufacturers “cherry-picking” the easiest and fastest option. It may be prudent, however, to eliminate the ability of manufacturers to select the Notified Body to which they submit their applications.

In the US, along with completing the reclassification process for devices on the market before 1976, the FDA should apply more stringent standards for acceptable predicates. Hines and colleagues (2010) discussed several issues with the existing use of predicates, including the permissive interpretation of intended use, disparate technological characteristics between the new device and predicate, and “predicate creep” (over time, a new device differs quite a bit from that of the original predicate). The agency has started to better clarify the use of predicates, which should help addresses some of these issues. Periodic audits of 510(k) applications and decisions may also help improve their adequacy, accuracy, and consistency.

Both jurisdictions could also encourage manufacturers to conduct pre-market studies. The current systems tend to reward “fast followers” to market that can take advantage of existing evidence from already marketed products. If eliminating the use of predicates in pre-market approval is not possible, fast followers could be required to generate the same clinical evidence as for other devices already on the market, unless there is compelling evidence of their comparable manufacture. Under such an approach, the first to market would set the evidence standard. This not only would reward manufacturers first to market by protecting against other manufacturers benefiting from their investment in clinical studies, but also would ensure that each new device is supported with evidence regarding its effectiveness, safety, and quality. Other actions to support the conduct of clinical trials and submission of quality clinical data are guidance on appropriate clinical trial designs to fulfil pre-market data requirements, new methods of streamlining clinical trials, and early scientific advice exchanged between the FDA and manufacturers.

When the evidence is insufficient at the time of approval, market access should be conditioned on rigorous prospective post-marketing studies to substantiate effectiveness and safety in real-world settings. Conditional approval would be one way to support innovation without burdensome overregulation while ensuring the patients' safety. Given manufacturer's poor record of completing such studies, US and European regulators should monitor studies more closely and take enforcement actions when they are delayed. In addition, comprehensive information on completed post-approval studies, including trial results, should be made easily accessible online. This would strengthen regulatory decisions and support "downstream" regulation by providing more robust evidence from which to inform pricing and reimbursement decisions. Ongoing investments in post-market data networks, such as Sentinel, electronic medical records, and UDI, may also help facilitate greater use of conditional approvals through better post-market data collection and analysis.

Post-market surveillance

Ensuring pro-active, not passive, post-marketing systems is just as important as strengthening pre-market authorisation. While reforms on the use of UDIs are a good step toward enabling the tracking and identification of devices, the true benefit of the UDI system will require its broad adoption and use by manufacturers, payers, providers, patients, and other stakeholders involved throughout the lifecycle of devices. Accordingly, we need strategies to facilitate the awareness, adoption, and implementation of the UDI system. Such efforts should focus on including UDIs in inventory logs, electronic health records, and claims data and linking different post-market databases, such as the GUDID and adverse event reporting repositories. Moreover, providers and patients should be engaged early to report, receive, and retain device information as well as to tailor strategies for communicating information (e.g. smart phone applications that can link the identifier to the UDI database) to different end users.

The UDI should be included in and facilitate the use of registries. Registries, which collect data on large numbers of patients using observational methods, may be a good way to monitor the use of devices in clinical practice and evaluate their long-

term safety and performance. Both Europe and the US have used registries to collect and evaluate post-market data, especially in cardiology and orthopaedics. For example, the National Joint Registries currently operate in the UK, Germany, and Italy, collecting information on hip, knee, and/or ankle replacement operations to monitor the devices' performance. In the US, the Kaiser Permanente Cardiac Device Registry tracks and monitors pacemakers and ICDs, with data on more than 22,000 ICD pulse generators, 52,000 pacemakers, and 90,000 leads. The registry allows the analysis of implant statistics, including complications, failures, replacements, usage, and costs.

Registries have been instrumental in identifying potential problems with a device or its use in practice (James et al. 2011). A recent analysis of the United Kingdom and French registries for transcatheter aortic valve implantation (TAVI) found that 25% and 20% of patients, respectively, were being treated transapically, which far exceeds what is justified by the clinical evidence and outside use approved by the FDA (Van Brabandt et al. 2012). The aforementioned SCAAR study on DES is another example (Lagerqvist et al. 2009).

Nonetheless, the use of registries needs to be improved. Because there is no consensus regarding which devices registries should include, we need criteria for when a device should be captured in a registry as a condition of approval and which devices might produce the most public health benefit from inclusion in registries. Ideally, this would also involve regulators working with stakeholders to establish basic standards for registries regarding methods, data quality, and transparency. Some of the main challenges with registries are adequately accounting for potential bias and the variability in the treatments, population, and settings captured, as well as the continuous development of devices. Creating a registries forum or consortium would be one way to bring the relevant parties together to share best practices and develop new methodologies for registries' data collection and analysis. New registries should also be linked to routinely collected health data, national mortality statistics, claims data, electronic health records, and other possible sources of relevant information. The implementation of UDIs should, in principle, increase the linkage of data.

Finally, given the continuous evolution of devices, regulators should require periodic update reports from manufacturers, especially for Class III devices. This could encompass information relevant to the device's benefits and risks, including new study results or scientific assessments of the device's risk-benefit ratio and estimates of the population exposed to the device. European regulators might stipulate a timeframe for beginning post-market device studies, similar to that of the FDA. Both jurisdictions should ensure that the results are publicly available in a timely manner.

Transparency of processes

Recent reforms, especially in Europe, have concentrated on improving transparency in device regulation. One particular focus of the proposals is improving the public's and health care professional's access to information. These stakeholders must have access to comprehensive information on the data submitted in the application (with due regard to commercial confidentiality when justified), the rationale for the Notified Body's decision, any post-market safety issues or defects, and devices that have been removed from the market. The EPARs used by the EMA may provide a model for communicating this type of summary. In addition, any request for information about a device not available publicly (from health care professionals, the public Competent Authorities, the commission, etc.) should be addressed without delay.

In the US, a public database of cleared devices would further aid transparency. The database could contain information about the device, a 510(k) summary, predicates used, and any other details pertinent to the clearance decision.

Table 7: Overview of existing weaknesses in US medical device regulation, recent reforms, and additional actions for improvement

Aspect of Regulation	Weaknesses	Recent Reforms*	Additional Actions Needed
Pre-market review	<ul style="list-style-type: none"> • Some high-risk devices undergo 510(k) route • Lack of or insufficient clinical data required/submitted for 510(k) and PMA applications • High/inappropriate use of predicates • Poor incentives for manufacturers to conduct new clinical studies • Lag in reclassifying “grandfathered” Class II devices • Inefficient <i>de novo</i> process 	<ul style="list-style-type: none"> • Modifications to the <i>de novo</i> process • New reclassification procedures • New guidances on 510(k) modifications, improving clinical trials, etc. • Investments in additional staff and industry training to enhance consistency and predictability of reviews 	<ul style="list-style-type: none"> • Apply more stringent standards for and review of acceptable predicates • Periodic audits of 510(k) applications and decisions • Require manufacturers that claim substantial equivalence to produce the same level of evidence as the first device to market • Guidance on appropriate clinical trial designs/data to fulfil pre-market data requirements • Identify and develop new methods of streamlining clinical trials • Early scientific advice between FDA and manufacturers • Greater use on conditional approvals tied to post-market studies, where appropriate
Post-market monitoring and surveillance	<ul style="list-style-type: none"> • Difficulty identifying and monitoring post-market device use • Poor rate of adverse event reporting • Subpar completion rate and quality of post-market studies 	<ul style="list-style-type: none"> • Develop UDI • Inclusion of devices in Sentinel surveillance system • Time requirement for submission of post-market study plans and initiation of studies. • Recalls system to identify devices most frequently subject of recalls and underlying causes of recalls 	<ul style="list-style-type: none"> • Include UDIs in existing databases, data networks, and health information technologies • Create strategies to engage provider and patient involvement in collecting, reporting, and retaining device information • Develop tailored communication strategies to relay device information to end users • Further development and use of registries • Require post-market device updates from manufacturers, especially for high-risk devices
Information provision on devices and regulatory processes	<ul style="list-style-type: none"> • Limited access to information contained in manufacturer applications and on major agency decisions, recalls, and adverse events • No publicly accessible information on unapproved applications • Limited transparency of clinical investigator financial interests 	<ul style="list-style-type: none"> • New guidances on agency review processes and requirements • Quarterly and annual reporting on progress toward meeting agency performance goals • Requirement for substantive summary of the rationale for any significant decision regarding a device application and review 	<ul style="list-style-type: none"> • Institute a public database of cleared devices with device and 510(k) information • Ensure results from post-market device studies are made publicly available in a timely manner

*Not an exhaustive list of reform actions

Source: Authors' compilation

Table 8: Overview of existing weaknesses in European medical device regulation, recent reforms, and additional actions for improvement

Aspect of Regulation	Weaknesses	Recent Reforms*	Additional Actions Needed
Pre-market review	<ul style="list-style-type: none"> • Devices generally only required to prove they are safe and work as intended (performance) • For most devices, limited requirements for clinical trial data; data submitted most often from laboratory tests, literature reviews, and small clinical studies • Lack of uniform standards and evidence requirements across Notified Bodies • Minimal coordination across Notified Bodies and Competent Authorities 	<ul style="list-style-type: none"> • Enhanced oversight and coordination of Notified Bodies • Greater harmonisation of the processes and evidence standards utilised by Notified Bodies • More centralised oversight of pre-market process by the European Commission and other expert groups • Certain high-risk devices must undergo review by specialised Notified Bodies • Notified Bodies must demonstrate sufficient in-house expertise to review applications 	<ul style="list-style-type: none"> • Require devices to demonstrate safety, performance, and effectiveness • More centralised review for all high-risk (and ideally medium-risk) devices • Consider eliminating ability of manufacturer to self-select Notified Bodies • Require manufacturers that claim equivalence to produce the same level of evidence as the first device to market • Guidance on appropriate clinical trial designs/data to fulfil pre-market data requirements • Identify and develop new methods of streamlining clinical trials • Early scientific advice between relevant scientific expert bodies and manufacturers • Greater use on conditional approvals tied to post-market studies, where appropriate
Post-market monitoring and surveillance	<ul style="list-style-type: none"> • Unclear and inconsistent adverse reporting to Eudamed • Vague guidelines on Eudamed requirements and use • Restricted information exchange opportunities via Eudamed • Difficulty identifying and monitoring post-market device use • No direct mechanism for health care professionals and patients to report device problems • Subpar completion rate and quality of post-market studies 	<ul style="list-style-type: none"> • Develop UDI • Required submission of more comprehensive information collected in Eudamed • Harmonised safety/adverse event reporting forms across member states • Regular submission of device safety reports by manufacturers • Extend Eudamed access to health professionals and the public • Institute deadlines for adverse event reporting 	<ul style="list-style-type: none"> • Include UDIs in existing databases, data networks, and health information technologies • Create strategies to engage provider and patient involvement in collecting, reporting, and retaining device information • Develop tailored communication strategies to relay device information to end users • Further development and use of registries • Require post-market device updates from manufacturers, especially for high-risk devices • Specific timelines for initiation of post-market studies
Information provision on devices and regulatory processes	<ul style="list-style-type: none"> • Limited access to information contained in manufacturer applications and on major Notified 	<ul style="list-style-type: none"> • Offer public access to Eudamed • Any publicly available information must be written in lay 	<ul style="list-style-type: none"> • Ensure stakeholder access to information on evidence submitted in CE application, rationale for Notified Body's decision, any post-market safety issues, and devices

	<p>Body decisions, recalls, and adverse events</p> <ul style="list-style-type: none"> • No publicly accessible information on unapproved applications • No EU-level repository of devices and relevant CE information • Post-market data not shared with the public 	<p>language</p> <ul style="list-style-type: none"> • Regular overviews of vigilance and surveillance information intended to be available to health care professionals and the public 	<p>removed from market</p> <ul style="list-style-type: none"> • Any requests for information should be made in a timely manner • Ensure results from post-market device studies are made publicly available in a timely manner
--	--	--	--

*Not an exhaustive list of reform actions

Source: Authors' compilation

Concluding remarks

Regulatory systems for medical devices have an important role in supporting market access to technological innovations, while duly protecting public health. In order to meet this aim, robust pre-market assessment and post-market vigilance are required. Both the US and Europe have recently introduced or are in the process of establishing reforms to meet this end. Such initiatives should be implemented in a timely manner, though additional actions will be required to enhance the reforms' effectiveness. More research is needed to assess the ongoing performance of regulatory approaches for devices.

SECTION II: HEALTH TECHNOLOGY COVERAGE, REIMBURSEMENT, AND PRICING POLICIES

Study 3: Comparative analysis of pharmaceutical coverage and pricing in Europe: Policy levers and mechanisms and insights for the United States

Introduction

Pharmaceuticals are a crucial and growing component of health care delivery. While increased demand for and use of pharmaceuticals has brought considerable benefits to patients and led to medical advancement, it has also resulted in burgeoning pharmaceutical spending (OECD 2010). This is the case in all developed countries, independent of the predominate source of expenditure and changes in the public-private mix of spending over time. In most countries, pharmaceutical expenditures are rising at rates greater than gross domestic product and, in some cases, more than other health care budgets (Vogler et al. 2008)⁵⁶. Consequently, rising pharmaceutical costs are considered an enduring challenge and one that requires a complex balancing act between expenditure control, affordable and equitable access to beneficial treatments, and support for innovation.

European governments have introduced a variety of policy approaches, particularly around regulating pharmaceutical pricing and coverage, to meet these often competing objectives. Such policies, and their perceived political and economic viability, vary across countries, reflecting distinct national traditions and policy priorities. Comparative drug evaluation, however, is increasingly employed in Europe to support evidence-based pricing and coverage decisions, as governments aim to attain better value for money from expenditures.

This paper undertakes a critical review of current policies across nine European countries - Denmark, England⁵⁷, France, Italy, Germany, the Netherlands, Sweden, and Switzerland. The paper first provides an analysis of national strategies to control pharmaceutical prices and set coverage and reimbursement levels. The discussion

⁵⁶ Comparing pharmaceutical expenditure levels across countries is beyond the scope of this paper. However, there are important differences in spending due to variations in price levels, national generic drug policies, therapeutic mix of drugs, volume and structure of consumption, and medical norms.

⁵⁷ The study focuses on pharmaceutical policy in England, but where indicated, some of the data presented refer to the UK as a whole.

then turns to the use of comparative drug evaluation in making pricing and coverage decisions. The paper concludes with a discussion of measures European countries have taken to support evidence-based coverage, reimbursement, and pricing decisions. Given the recent investment in CER in the US and drive toward obtaining greater value for the high levels of health care spending (Antos et al. 2013; Ginsburg et al. 2012; Schoen et al. 2013), the paper also discuss some lessons American policy makers might consider based on the European experience.

National policy measures to control pharmaceutical prices and set coverage levels

To regulate access to pharmaceuticals and control prices, European countries employ a range of policy measures (Table 8), with the majority aiming to address supply-side objectives. In particular, these tools aim to ensure efficient spending, equitable and clinically appropriate access to high-quality medicines, and, in some countries, a vibrant domestic pharmaceutical industry (Mossialos et al. 2004).

Countries apply direct price controls to on-patent drugs, except in Denmark, England, and Germany, where prices can be freely set by manufacturers or, in some instances, wholesalers at market launch. However, prices are indirectly influenced by the reimbursement system in Denmark and Germany and within the context of the Pharmaceutical Price Regulation Scheme (PPRS) in England. In the latter case, indirect controls, principally in the form of profit or rate of return regulation, aim to ensure that manufacturers do not receive excessive profits on products financed by the National Health Service (NHS), while concurrently rewarding innovation (Mossialos et al. 2004).

Within the framework of price controls, countries use a range of strategies to set price limits or define reimbursement amounts (Table 8). Statutory pricing, the most commonly used method, involves comparing proposed prices for new products against those prices paid by other payers (external reference pricing or international price comparisons), or against those prices already paid for products judged to be similar (internal reference pricing).

External reference pricing is used in all countries to set prices, except those that free price and Sweden⁵⁸, typically based on comparisons using a basket of reference countries, selected according to geographic proximity and economic comparability. Germany, England, and France are the three countries most commonly referenced, driven largely by the fact that they often have prices available for reference soon after launch. The prices in these countries are commonly used informally in price negotiations internationally, even if they are not used in formal reference price schemes. Countries also differ in the way in which international benchmark prices are used, including the type of drug covered (e.g. reimbursed, patented), at what point in time comparisons are made, and where in the supply chain prices are derived.

In practice, most countries use external reference pricing as one of a range of measures to set prices or only apply it to a limited range of drugs (Mossialos et al. 2006). For example, France uses this approach only for reimbursed, highly innovative products. This is likely due to the technical challenges in making international price comparisons, such as the sensitivity to sample selection and units of measurement, as well as the general difficulty in establishing what prices are actually paid in reference countries given the extensive use of rebates and other pricing mechanisms (Danzon and Kim 1998; Mossialos et al. 2006). Moreover, such comparisons can create access issues, as manufacturers may be incentivised to delay launch in lower-priced markets (OECD 2008).

Internal reference pricing is used in Denmark, Germany, Italy, and the Netherlands to define the maximum price or reimbursement level for defined groups of interchangeable drugs, typically within a therapeutic class⁵⁹. These groups of drugs have generally included generic drugs only, but the Netherlands includes on-patent drugs and Germany has recently started to do so. In countries using this approach to set reimbursement limits, as opposed to a direct tool for price regulation

⁵⁸ However, Sweden plans to introduce external reference pricing in 2014 and Germany recently established, via the new 2011 German Act to Reform the Market for Medicinal Products (AMNOG), a process where if pricing negotiation fail between the GBA and manufacturers, an arbitration board sets the price on the basis of the comparator's price, a GBA assessment, and the price in 15 other European countries.

⁵⁹ Drugs are normally grouped together in the same reference group if they have the same chemical ingredients and have comparable therapeutic effects based upon the concept that they are interchangeable.

(Germany, Denmark, and the Netherlands), coverage levels are often set at the price of the cheapest drug (often generic) and most drug prices converge at the reference price, as there are limited financial incentives to price below the reference price (Danzon and Ketcham 2003; Kanavos and Reinhardt 2003; Lopez-Casasnovas and Puig-Junoy 2000). For drugs priced above the reference price, payers either fund part of the premium price in the form of higher reimbursement or the additional cost is shifted to the patient through increased cost-sharing, with the latter being more common in most cases.

In principle, reference pricing attempts to stimulate demand-side cost containment by shifting the difference in prices between branded and generic substitutes to the consumer, thereby incentivising patients and physicians to consider drug prices in decision making. However, the European experience suggests that this is rarely realised, as manufacturers adapt their prices to coincide with the reference price and physicians may prefer to prescribe products that are not included in the reference price scheme to avoid discussing co-payments with patients. Similar to external reference pricing, this approach introduces additional technical challenges, in that implicit judgments are often required for valuing differences across drugs as well as a robust empirical methodology (i.e. how products are clustered and reimbursement amounts set). This approach has gained popularity, however, because it can be effective in reducing price differences among drugs defined as therapeutic substitutes by improving market transparency (Giuliani et al. 2003).

Internal reference pricing is also used in France, Italy, and Sweden⁶⁰ to regulate the price of generic drugs. The generic is priced at market entry at a discount by reference to the price of the original product; the amount of the discount is specified by the regulator. In France, generic drugs must be priced at least 50 percent below the originator price. Other countries require a discount ranging from 20 percent (Italy) to 40 percent (the Netherlands) (OECD 2010). Increasingly, regulators also revise the price of originals at market entry; this is the case of France, where price reductions are not only suggested for the originator product, but also for other on-patent drugs of the same therapeutic class at the time of generic listing.

⁶⁰Technically more of a system of obligatory generic substitution, where substitutable drugs are clustered and prices not exceeding the highest price within such a group are automatically accepted for reimbursement.

As illustrated in Table 9, several countries apply both external and internal reference pricing in pricing and/or coverage determinations. This dual-approach functions effectively as a double budgetary “safety net”; external reference pricing provides a safeguard that prices do not exceed those elsewhere, where internal reference pricing protects against significant price differences within therapeutically equivalent medicines.

One strategy increasingly used in Europe is CER or, more broadly, HTA⁶¹ to inform decisions about which drugs to include on national lists or formularies⁶², reimbursement levels, and, in some cases, pricing. It is also used to encourage high quality care by identifying which patients are most likely to benefit from treatment, optimal patterns of use, and appropriate placement in the spectrum of care (Drummond 2003). This approach is particularly well established in England, the Netherlands, and Sweden, although it is applied differently across countries, as discussed in further detail in the following sections.

Finally, countries also use a number of ex-post mechanisms⁶³, which seek to minimise the financial impact of a drug after its initial entry onto the market (Mossialos et al. 2003). These include price-volume agreements (France, Germany, Italy, Sweden); price or margin cuts (Denmark, England, France, Germany, Italy, Sweden, Switzerland); payback or clawback policies (England, France, Italy); and, discounts or rebates (France, Germany, Italy). Many of these mechanisms are considered blunt policy instruments and therefore may have unintended consequences on other policy measures. For example, payback policies may reduce

⁶¹ HTA normally includes CER and health economic evaluation, a form of comparative analysis that, unlike CER, incorporates costs (and, often other factors such as social, ethical and legal considerations) as well as health outcomes.

⁶² Almost all countries use a ‘positive’ list to indicate which drugs are eligible for public reimbursement (covered) in part or in full. The exceptions are England and Germany, which use a national ‘negative list’ to delineate products excluded from coverage. However, in these countries regional health authorities may formulate their own positive lists. Some countries have more than one positive or negative list due to different eligibility criteria and/or reimbursement rates. France, for example, has separate lists of innovative drugs, allowing for special funding arrangements. Spain has both positive and negative lists.

⁶³ *Price-volume agreements*: prices are set according to expected or realized volume, such that if volume passes a threshold, the price level will decrease and/or companies have to repay the government or associated health insurance plan; *price or margin cuts*: Prices are cut (or frozen) by law or as an outcome of a negotiated agreement; *paybacks/clawbacks*: requires manufacturers to pay back a share of their revenue, if a pre-specified budget ceiling for public pharmaceutical expenditures is exceeded or applies to pharmacies (clawbacks), requiring them to pass a part of their turnover to third party payers; *discounts/rebates*: imposed on manufacturers and pharmacists, such that they have to return a part of their revenue.

price transparency, as it changes the effective price, but not the list price, thereby reducing the effectiveness of external reference pricing. However, these instruments have been largely effective in controlling expenditure (Espin and Rovira 2007). In fact, some of these policies have been introduced to attain short-term desired savings during the recent economic crisis (Vogler 2011).

Table 9: *Strategies to control outpatient drug prices and set coverage and reimbursement levels*

Strategies	CH	DK	EN	FR	DE	IT	NL	SE
Branded drugs								
Free pricing		Prices freely set by manufacturers, but indirectly influenced by the reimbursement system; obligatory price notification to Medicines Agency	Prices freely set by manufacturers, but indirectly controlled through PPRS		Prices freely set by manufacturers, but indirectly influenced by the reimbursement system; obligatory price notification to Pharmacists Association However, external-reference pricing-like procedures introduced in new 2012 AMNOG law			
Profit controls			Free pricing subject to rate of return regulation (PPRS)					
External reference pricing	Applies to reimbursable drugs; references DK, DE, NL, UK (AT, FR, IT may also be considered)			Applies to reimbursable, innovative drugs; references DE, IT, SP, UK		Only used as additional information during price negotiations for reimbursable drugs; references all EU member states	Applies to prescription-only medicines; references BE, DE, FR, UK	Will begin to use in 2014
Internal reference pricing		Applied to set reimbursement limits (not in price regulation); reference groups include generics only (via generic substitution scheme)			Applied to set reimbursement limits (not in price regulation); reference groups include generic and branded drugs	Only used during price negotiations to determine reimbursement decision for reimbursable drugs	Applied to set reimbursement limits (not in price regulation); reference groups include generic and branded drugs	

Price-volume agreements				Two general types of schemes used: 1) payback mechanism for excessive sales by therapeutic class, based on manufacturer's agreed turnover, and 2) special rebates for certain drugs where prescribing volumes in France are high compared to other countries	May be agreed during price negotiations between manufacturer and the payer group; may be part of a risk-sharing agreement	For a limited number of pharmaceuticals (e.g. expensive drugs); agreement negotiated during the pricing procedure. Often entails a payback clause, where manufacturers must pay back a proportion of agreed target budgets (to the Ministry of Health) if they sell more drugs than expected		Typically used in the case of innovative drugs as part of risk-sharing agreement
Price or margin cuts	Regular price reviews to assess price revisions	Agreement on reduction in overall price level such that overall expenditure on subsidized pharmaceuticals is kept constant	Price cut of 1.9% on branded NHS medicines as part of 2009 PPRS	Periodic price reductions for new and expensive products	Price freeze on reimbursable medicines (2010) and changes in structure to whole margin from 2012	Frequent use of price cuts and wholesale margin cuts		Annual price reviews to assess price revisions
Paybacks/clawbacks			Linear statutory wholesale mark up with clawback	A predetermined turnover rate is set for each producer Three thresholds ranging from 50-70% payback growing with excess consumption Innovations, generics, and orphan drugs are exempt		Industry pays 40% excess consumption		

Discounts/rebates				Rebates calculated for each manufacturer based on drug's innovativeness (added therapeutic benefit and share of the increase in expenditure) Orphan drugs are exempt	Mandatory manufacturer's rebate to social health insurance (originally 6% and increased to 16% in 2010)	Choice between payback and price cuts		
HTA	No formal HTA requirements, but submission of health economic information by manufacturer recommended	No formal HTA requirements, but submission of health economic information by manufacturer recommended	National Institute of Health and Clinical Excellence (NICE) – assessments inform coverage decisions for select drugs (e.g. high health or budget impact)	No formal HTA requirements, but submission of health economic information by manufacturer recommended	Institute for Quality and Efficiency in Health Care (IQWiG) – assessments inform coverage decisions/reimbursement price for drugs outside of reference pricing system	Italian Medicines Agency (AIFA) – assessments only required to inform pricing decisions for innovative drugs	Health Care Insurance Board, Committee for Pharmaceutical Aid (CHF) – assessments inform coverage/reimbursement price for drugs outside of reference pricing system	Dental and Pharmaceutical Benefits Board (TLV) – assessments inform pricing and coverage decisions for drugs
Generic drugs								
Free pricing		Prices freely set by manufacturers; but indirectly influenced by the reimbursement system; obligatory price notification to Medicines Agency						
Internal reference pricing	Priced at least 30% lower than price of original product			Priced at least 50% lower than price of original product		Priced at least 20% lower than price of original product	Priced at least 40% lower than price of original product	Priced lower or equivalent as the highest price within a group of substitutable products

Sources: Carone et al. (2012), IMS (2010), Vogler et al. (2008), Vogler et al. (2011), Vogler (2012)
Denmark (DK), England (EN), France (FR), Germany (DE), Italy (IT), the Netherlands (NL), Sweden (SE), Switzerland (CH)

Determining pharmaceutical pricing and coverage

Given some of the limitations of more traditional direct control policies (e.g. price controls), decision makers have placed greater emphasis on making explicit assessments of a drug's therapeutic benefit and, in some cases, cost-effectiveness to inform coverage and pricing determinations. Such an approach increases the scope to obtain better value from expenditures, where decision making is based on evidence of value (not simply cost) and, in principle, grounded in greater transparency. The following sections discuss how pricing and coverage decisions are governed and organised, and the processes and methods employed by different countries to make such determinations (Table 10).

General framework and governance

Across most countries, pharmaceutical coverage and pricing entails a multi-staged process. A *review or assessment* of a drug's benefits, relative benefits, and costs is first undertaken, followed by an *appraisal* (i.e. interpretation and consideration) of the evidence generated or considered in the assessment phase to inform the coverage and, sometimes, pricing decision. To that end, pricing and coverage determinations generally entail separate processes, but there are some synergies between these two functions in each country⁶⁴. Coverage recommendations or decisions directly influence pricing determinations (to varying degrees) in some countries, namely France, Germany, the Netherlands, Sweden, and Switzerland. In Germany, for example, evidence from the benefit assessment is used in price negotiations between the GBA and payers. In all the other jurisdictions, the relationship between coverage and pricing determinations is more implicit and indirect. Most countries consider the issue of 'value' at the coverage stage to help determine whether a product should be covered and under what conditions. Such considerations are not typically used to directly set the price of a treatment, but rather to determine the reimbursement rate and to substantiate whether it will receive a premium price over other drugs.

⁶⁴Pharmaceutical pricing and reimbursement generally follows a two-step process: first, a drug is accepted for coverage and second, its price is determined. In practice, however, except in countries that have free pricing, a loose agreement on price is reached before the coverage decision is made and a final price is agreed afterwards.

Centralised national agencies are involved in pharmaceutical pricing and coverage⁶⁵; these two functions may or may not be carried out by the same body. In the European countries examined, the agencies assume different roles, in terms of decision-making authority and relation to government (Hutton et al. 2006; Sorenson et al. 2008a). Some agencies act in a regulatory capacity, making decisions about coverage and/or pricing. In several cases, these bodies typically possess an arms-length relationship to government. Others take an advisory role, making coverage and/or pricing recommendations to government, often the Ministry of Health, which then a final determination. In countries with a direct relationship with government (e.g. France, the Netherlands), the Ministry of Health oversees the assessment process or set priorities for assessment to some degree (Hutton et al. 2006). The entities involved in pharmaceutical coverage and pricing can also be categorised as those that “produce” evidence (England, Sweden) – that is, they conduct evidence synthesis, economic modelling, and other studies – and those that mainly “use” existing evidence, typically submitted from manufacturers, to support coverage and pricing recommendations or decisions (Denmark, France, Italy, the Netherlands, Switzerland).

Prioritising reviews

Countries use different mechanisms to determine which drugs to review. Denmark, France, Italy, Sweden, and Switzerland, for example, evaluate every new drug before making a decision about coverage⁶⁶. In England, only drugs referred to NICE are candidates for review, which are then prioritised using a variety of criteria, including health impact, disease burden, and clinical and policy relevance. In comparison, the Netherlands only assesses drugs that cannot be classified by existing reference pricing systems, while Germany limited evaluation to drugs that have a new active ingredient (or combination thereof), have economic significance to the SHI system, and are used in multiple setting (not only in hospitals).

⁶⁵ See Velasco-Garrido et al. (2008) for a comprehensive overview of other national HTA organisations and activities.

⁶⁶This entails reviewing every new drug dossier submitted by manufacturers to support a coverage determination. Thus, in principal, manufacturers ultimately decide which drugs are reviewed.

Regardless of the prioritisation approach adopted, countries aim to ensure that the prioritisation process is as sound and transparent as possible by involving a variety of stakeholders in the process, publishing prioritisation criteria, and/or soliciting public opinions on social values and priorities. However, the degree of openness and transparency of the process does vary across countries. England is arguably the most inclusive, allowing a range of stakeholders (e.g. health experts, Department of Health, patient groups) to recommend potential drugs for review. NICE in England also publishes its criteria on its website and regularly gathers public perspectives on review priorities and important coverage decision criteria.

Assessment requirements and procedures

There are distinct differences across countries regarding what type of evidence or methods are required and the processes involved in the assessment. Such differences are due largely to differences in resource and regulatory constraints, as well as health system objectives. Several countries publish guidelines outlining their evidence and methodological requirements, but the guidelines often vary in detail and transparency (Sorenson et al. 2008a; Sorenson 2010).

All countries require a clinical effectiveness assessment (i.e. therapeutic benefit and safety). In most cases, the assessment is based on evidence supplied by the manufacturer, but may also be informed by research conducted in-house or by an independent group, including systematic reviews and meta-analyses of available data. Data from RCTs are generally preferred to assess a drug's comparative effectiveness, while a variety of approaches (cost-effectiveness, cost-utility, and cost-benefit analyses) are accepted to demonstrate economic costs and benefits. Data to substantiate economic value typically derive from different types of data (clinical, epidemiological, demographic, economic) from different sources (studies, registries, databases, models). For countries requiring cost-utility analysis, a quality-adjusted life year (QALY) is used to measure health benefit (and cost per QALY to ascertain value for money). Although the QALY is broadly accepted, there is continuing debate regarding its limitations (Oliver and Sorenson 2010). As an alternative, countries such as France, Germany, Italy, and Switzerland have largely rejected the QALY approach and instead favor the use of disease-specific measures.

Only England, the Netherlands, and Sweden require an economic assessment in drug reviews. In Germany, a cost-benefit analysis is only required if no additional benefit or therapeutic improvement has been found to exist and the manufacturer and Federal Association of Statutory Health Insurance Funds cannot reach agreement regarding price. Recent reforms in France (decree 2012-1116) will result in a requirement for economic assessments in select circumstances, namely when the claimed therapeutic benefit of the drug falls within the major, important, or moderate categories and the drug is likely to have a significant impact on care organisation, clinical practice, patient care, or costs (Remuzat et al. 2013). Such requirements will be extended to drugs applying for renewal of inclusion on the reimbursable drugs formulary.

The data available for assessments, however, often are not sufficient or conclusive. Head-to-head RCTs, which are preferred by assessors, are generally lacking, particularly prior to or by product launch. This is partly influenced by the fact that the EMA does not require a comparative assessment of a drug's efficacy, safety, and quality with existing therapies in pre-marketing licensing (Eichler et al. 2010; Sorenson et al. 2011b). Beyond problems of quantity, there are issues with data quality, namely that studies do not always compare all possible treatments or patient populations and often fail to measure all relevant outcomes, particular economic data.

To address these issues, some countries (England, Italy) offer product-specific scientific advice to manufacturers by reviewing early product development plans to assess whether they will generate the relevant evidence for future assessment. In addition, HTA bodies or payers in England, France, Italy, the Netherlands, and Sweden use risk-sharing agreements and CED schemes in limited circumstances to allow temporary coverage and reimbursement while the outcomes of a drug are being further substantiated. Risk-sharing agreements, for example, allow coverage based on meeting certain, specified conditions, such as cost, volume, market share, and cost-effectiveness targets (Adamski et al. 2010; Cook et al. 2008). If the conditions are not met, then coverage may be withdrawn or the drug's price reduced. For example, after NICE controversially recommended against the use of various

products for multiple sclerosis, the government established a risk-sharing scheme with manufacturers to supply these treatments on the NHS. Under the scheme, patients were monitored annually and the amount paid for the drugs was adjusted on a sliding scale if patient outcomes differed from an agreed cost per QALY of no more than £36,000 (\$59,000). The CED approach applies a similar strategy. Coverage is condition, based on the collection of post-market evidence and re-evaluation (Hutton et al. 2007). These types of approaches are particularly suitable for severe conditions or areas of high unmet need, high-cost drugs, and situations where there is strong political or patient lobbying for access (Sorenson et al. 2010).

In addition to its CED program, NICE in England is piloting an “Innovation Pass” scheme in collaboration with the Departments of Health and Science and Innovation, with the aim to help patients with rarer disease to get innovative new drugs not yet appraised by the institute. Select patients will be given access to the drug while additional data collected, which will contribute towards a future NICE assessment and appraisal.

Evidence appraisal and coverage and pricing determination

Based on an appraisal of the evidence, a decision is taken on whether to grant coverage and, in some cases, to ascertain the level of reimbursement and a final price. In each of these countries, a drug’s relative therapeutic benefit is the most important criterion in determining coverage status, followed by cost-effectiveness, where applicable (Sorenson et al. 2008a). Other criteria considered in coverage decisions include patient benefit (health-related quality of life); disease severity; availability of alternative therapies; public health impact; degree of innovation; budget impact; and social and ethical aspects, such as equity.

Cost-effectiveness is particularly important for drugs that have new indications, are expensive, are expected to be widely used, or whose benefits differ by indication or patient sub-group. Some countries (England, the Netherlands, Sweden), which explicitly use cost-effectiveness in coverage decision making, use a cost-effectiveness or price threshold to establish whether a drug provides sufficient value and to determine coverage status and, in some cases, reimbursement levels. A

threshold generally represents the amount of money a society is willing to pay for an additional unit of health outcome (i.e. an additional QALY). Such “decision rules” are often implicit and case-dependent. The value of the cost threshold varies by county: it is generally set at £20,000-£30,000 (\$30,000-\$45,000) in England; €20,000 (\$30,000) in the Netherlands; and 500,000SEK (\$62,000) in Sweden (Devlin and Parkin 2004; Zwart-van Rijkom et al.; Perrson and Hjelmgren 2003). The use of a threshold and at what value(s) it is (or should be) set at is an issue of considerable debate, with concerns that it is arbitrary, does not capture important measures of value, and may not reflect societal preferences (Kennedy 2009; McCabe et al. 2008). Several countries have formally taken into account these issues. For example, the Netherlands and Sweden are considering adopting a revised approach that adjusts the threshold according to need, severity of disease, or equity considerations, especially for certain high-cost drugs (e.g. orphan and cancer drugs). Following public consultation, England recently agreed to extend its threshold for drugs aimed at end of life care under some circumstances to facilitate cancer drug access in the NHS.

Instead of a threshold France and Germany employ a ranking measure to represent therapeutic value and the relative benefit for new drugs. In France, the therapeutic value of a drug is ranked across five categories, including major benefit, important benefit, moderate benefit, weak benefit, and insufficient benefit⁶⁷. Similarly, Germany’s ranking categorises drugs as follows: major added benefit, significant added benefit, small added benefit, an additional benefit, no added benefit, and benefit below alternative therapy(s). The main difference between these two ranking systems is that France only assesses absolute benefit, while Germany evaluates comparative value between the new drug and a specified comparator(s). However, France does consider relative benefit in its pricing determinations.

Although countries are becoming more selective in coverage determinations, especially for expensive products, it is rare for an approved drug not be accepted for

⁶⁷ France is planning to implement a new system called the Relative Therapeutic Index (ITR), which will combine the ranking scales currently used for reimbursement and pricing determination. Evidence of therapeutic benefit of a new drug will be assessed against a relevant comparator(s) and ranked accordingly: -1 (inferiority compared to the relevant comparator or use of a non-relevant comparator, unacceptable methodology, or lack of evidence), 0 (non-inferiority compared to the relevant comparator), 1 (minor improvement compared to the relevant comparator or improvement of conditions of use with impact of care and non-inferiority compared to the comparator), 2 (moderate improvement compared to the relevant comparator), or > 3 (major improvement compared to the relevant comparator).

any level of coverage. For example, over the last ten years, NICE, who arguably employs the most comprehensive reviews of the countries examined, only gave a negative recommendation to around seven out of its 110 or so drug appraisals. However, it is common for drugs to be covered with conditions based on use in certain indications, patient groups, treatment settings, or consumption levels. For instance, Germany and Sweden recommend specific first-line treatments for coverage, while the Netherlands determines coverage based upon patient population, indication, and prescribing physician (Sorenson et al. 2008a; Vogler et al. 2008).

In addition to any conditions to coverage, positive coverage decisions generally outline the extent of reimbursement and patient cost-sharing (if any) and any exemptions. All of the countries except the Netherlands employ some form of cost-sharing for outpatient prescription drugs in the form of coinsurance (France), deductibles with coinsurance (Denmark, Germany, Sweden, Switzerland), or co-payments (England, Italy). France is the only country where the level of cost-sharing coincides with a drug's assessed effectiveness (Remuzat et al. 2013). Cost-sharing arrangements range from 30% (changed from 35% in May 2011), 65%, and 100%, based on “major” or “important”, “moderate”, and “weak” or “insufficient” benefit, respectively⁶⁸. This tiered payment system aims to motivate patients (and their physicians) to choose high value drugs. Drugs deemed as irreplaceable and particularly expensive (e.g. HIV drugs) are covered in full. In countries that reference price, patient are often required to pay any price above the maximum reimbursement price. This approach assumes that some patients will be willing to pay for the additional benefits provided by higher-price (generally newer) drugs. Consumers then, in turn, send signals regarding the value they place on certain benefits. In practice, however, it is unclear that patients have the necessary information and ability to ascertain the relative benefit across products in a meaningful way. In all countries, however, positive lists are fairly comprehensive and cost-sharing is usually low, allowing patient access to medically-needed therapies⁶⁹. Moreover, patients falling under certain categories (e.g., children under 18 years, chronically ill, low income) are regularly exempt (Thomson and Mossialos 2010).

⁶⁸ Once the ITR (see footnote 55) is implemented, the level of cost-sharing will depend on which one of the five categories the drug is ranked and the level of reimbursement of the comparator used.

⁶⁹ If a drug is not publicly covered, it may be available through (voluntary) private health insurance, which is the case in England, the Netherlands, possibly Sweden, and Switzerland. However, individuals may be more likely to pay out-of-pocket for unlisted drugs.

Some countries use evidence of value to also directly inform pricing decisions⁷⁰. In Sweden, manufacturers are required to apply to the TLV for coverage and reimbursement at a proposed price, so these decisions are taken concurrently based on clinical and cost-effectiveness evidence. While the TLV does not negotiate the price of the drug based on the evidence, pricing may be varied by patient subgroup (i.e. use in some groups may be more cost-effective than others). There were plans to adopt a similar approach in England, but it appears this will no longer be the case and the system of free pricing will remain in place. In France and Germany, the use of evidence is more indirect, where a drug's therapeutic advantage over existing alternatives is considered in price negotiations with manufacturers.

According to a European Commission Directive (Transparency Directive 89/105/EEG), the coverage, reimbursement and pricing process is required to be completed within 180 days of application submission. In most countries, however, it takes longer to get to a final decision(s). The review process alone takes three months to two years on average across countries; longer review times are normally taken for cancer drugs and other complex therapies. The time to decision is a frequent point of contention with policy makers, industry, and patient groups. France, the Netherlands, and Switzerland offer expedited review processes for particularly innovative drugs or those that treat life-threatening illness, and England uses a single technology appraisal (STA) process to limit assessments to a single technology. STAs place more emphasis on manufacturer data and less on extensive systematic review and expert consultation, compared to its usual multiple technology appraisal (MTA) process.

Ex-post re-evaluation of pricing and coverage status

While the drug assessment and appraisal process typically occurs prior to market launch, some countries (England, France, the Netherlands, Sweden, Switzerland) also undertake systematic re-evaluations after drugs have been used in

⁷⁰ In some countries, particularly those with reference pricing systems, evidence of therapeutic benefit can indirectly influence pricing decisions to the extent that it helps form a judgement about whether or not a drug offers additional therapeutic value relative to similar products (ie, reference groups). If so, it may be granted a higher price than the reference amount.

practice to identify products that do not demonstrate good value or those that have become obsolete. Re-evaluating allows a greater range of drugs to be assessed for value, especially in countries where not every drug is reviewed, and helps ensure optimal resource use and patient care.

Evidence from post-market reviews can be used to modify pricing and coverage status, where appropriate, or to determine areas of disinvestment (i.e. removal from list of publicly covered drugs). For example, Denmark operates a five-year review process of the pricing and coverage status of existing drugs, and the TLV in Sweden has been evaluating all drugs approved prior to 2002 (Vogler 2008). In the past ten years, NICE in England has identified over 800 interventions for disinvestment (Garner and Littlejohns 2012) – an activity supported strongly by policy makers in recent years (Darzi 2008).

Table 10: Comparative approaches to pharmaceutical pricing and coverage decisions

	CH	DK	EN	FR	DE	IT	NL	SE
Structure and Organisation								
<i>Key agencies</i>	Federal Office of Public Health (FOPH) and the Federal Drug Commission (FDC)	Reimbursement Committee of the Danish Medicines Agency (DKMA)	National Institute for Health and Clinical Excellence (NICE)	National Health Authority (HAS): Transparency Commission (coverage); Economic Committee for Health Products (pricing)	Federal Joint Committee (GBA) and Institute for Quality and Efficiency in Health Care (IQWiG)	Italian Medicines Agency (AIFA): Scientific and Technical Committee (coverage); Pricing and Reimbursement Committee (pricing)	Health Care Insurance Board, Committee for Pharmaceutical Aid (CHF)	Dental and Pharmaceutical Benefits Agency (TLV)
<i>Role in coverage and reimbursement decisions</i>	FDC evaluates and classifies new drugs and makes recommendation to FOPH, who renders a final coverage decision	DKMA committee reviews evidence and renders coverage decision	NICE reviews evidence and makes renders coverage decision	Transparency Commission reviews evidence and advises on coverage decisions; Ministry of Health makes final listing decision; Economic Committee for Health Products considers evidence and renders pricing decision	GBA reviews evidence and may commission the IQWiG to conduct a benefit assessment. GBA produces report, consults with stakeholders, and renders final coverage decision (initial benefit determination). Price negotiation and arbitration transpires between manufacturer and the Federal Association of Statutory Health Insurance Funds.	Reviews evidence and renders coverage and pricing decisions	CHF reviews evidence and makes recommendations on coverage and pricing; Ministry of Health, Welfare, and Sport makes final listing and pricing decisions	Reviews evidence and renders coverage and pricing decisions
<i>Relationship to government</i>	Integrated	Integrated	Arms-length	Integrated	Integrated	Arms-length	Integrated	Arms-length
Requirements and Methods								
<i>Prioritisation criteria</i>	Every new drug ^a	Every new drug ^a	Drugs referred by Department of Health (other stakeholder can also nominate	Every new drug ^a	Every new drug that have new active ingredients or new combinations of active ingredients.	Every new drug ^a	Drugs that cannot be classified within reference pricing system	Every new drug ^a

			drugs for review), which are then prioritized based on a variety of criteria, such as health impact, disease burden, and clinical/policy relevance.		Drugs that are of limited economic significance (<1 million Euro/yr with SHI) and those only used in hospitals are excluded for benefit assessment.			
<i>Evidence requirements</i>	<p>RCT data for clinical efficacy preferred; health economic, epidemiologic, and disease burden information accepted, if available; comparative effectiveness with existing therapies (via internal reference pricing or therapeutic benefit); summary of the three most relevant clinical papers; Swiss and foreign physicians drug prescription information; request for and justification of an innovation bonus.</p> <p>Source: evidence from manufacturer dossier</p>	<p>RCT data for clinical efficacy preferred; health economic information recommended, but not required to establish value for money (usually submitted to justify a high price)</p> <p>Source: evidence from manufacturer dossier.</p>	<p>RCT data for clinical efficacy preferred; health economic information required to establish value for money. Other evidence on societal preferences, equity impacts, innovative characteristics, and budget impact may be submitted.</p> <p>Source: systematic reviews and analyses of clinical and economic studies; may or may not include manufacturer data.</p>	<p>RCT data for clinical efficacy preferred; health economic information recommended, but only required in certain circumstances to establish value for money. Other evidence on public health impact, innovative characteristics, and budget impact may be submitted.</p> <p>An economic assessment is only required when the benefit claimed by the manufacturer fall in the major, important, or moderate categories and the drug is likely to have a significant impact on care organisation, clinical practice or patient care. The price of the drug may also warrant an economic assessment (if high).</p> <p>Source: evidence from manufacturer dossier</p>	<p>RCT data for clinical efficacy preferred; number of patients and patients groups for which an additional benefit is claimed; therapy costs for SHI; and, safety information.</p> <p>A cost-benefit evaluation is only considered after an arbitration award if no additional benefit or no therapeutic improvement has been found to exist and the manufacturer and Federal Association of Statutory Health Insurance Funds cannot reach agreement regarding price.</p> <p>Source: evidence from manufacturer dossier</p>	<p>RCT data for clinical efficacy preferred; health economic information recommended, but not required to establish value for money (mostly for pricing purposes); budget impact; estimated market share; prices and consumption data in European countries.</p>	<p>RCT data for clinical efficacy preferred; health economic information required to establish value for money. Other evidence on innovative characteristics and budget impact may be submitted.</p> <p>Source: evidence from manufacturer dossier</p>	<p>RCT data for clinical efficacy preferred; health economic information required to establish value for money. Other evidence on disease burden/severity and equity impacts may be submitted.</p> <p>Source: systematic reviews and analyses of clinical and economic studies; may or may not include manufacturer data.</p>

<i>Assessment of health benefit</i>	Comparative benefit assessment categorization	Not specified	QALY	Comparative benefit assessment categorization If economic assessment required, LY or QALY used depending on the type of assessment	Comparative benefit assessment categorization	Comparative benefit assessment categorization	QALY	QALY
<i>Assessment of economic value^b</i>	CEA	CEA, CUA	CUA (CEA considered in some cases)	CEA, CUA	CBA (but only in certain circumstances)	CEA, CBA	CUA (CEA considered in some cases)	CUA (CEA and CMA considered in some cases)
<i>Choice of comparator</i>	Existing treatment alternatives	Not specified	Current best alternative or routine treatment	All relevant competing interventions (current best or routine treatment most common)	Selected by GBA on case-by-case basis	Most widely used treatment	Routine treatment	Three comparators required from same therapeutic group: 1) routine treatment, 2) nonmedical intervention, and 3) no treatment
<i>Principal outcome measures</i>	Effectiveness, appropriateness, and cost-effectiveness	Relevant to specific treatment, willingness to pay (but only to supplement main outcomes)	Mortality, morbidity, quality of life	Mortality, morbidity, quality of life	Mortality, morbidity, health-related quality of life	Mortality, morbidity, disease specific endpoints	Mortality, morbidity, quality of life	Mortality, morbidity, quality of life, willingness to pay
<i>Costs</i>	Direct costs (mainly treatment costs)	Direct costs; if indirect costs are included, must be reported separately	Direct costs; indirect costs, depending on the assessment	Varies; if indirect costs are included, must be reported separately	Direct costs; indirect costs	Direct and indirect costs	Direct costs; if indirect costs are included, must be reported separately	Direct costs; indirect costs, depending on the assessment
<i>Sensitivity analysis</i>	Not available	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Subgroup analysis^c</i>	Not available	Not specified	Yes	Yes	Yes	Yes	Yes	Yes
Applications to Coverage and Reimbursement Decisions								
<i>Applications of evidence</i>	Used to make decisions on inclusion in the benefit schedule and any pricing or reimbursement conditions.	Used to make decisions on inclusion in the benefit schedule and any pricing or reimbursement conditions.	Used to make decisions on inclusion in the benefit schedule and any pricing or reimbursement conditions. Also	Used to make decisions on inclusion in the benefit schedule and any pricing or reimbursement conditions.	Used to make decisions on inclusion in the benefit schedule and final pricing decision.	Used to make decisions on inclusion in the benefit schedule and reimbursement conditions. Also	Used to make decisions on inclusion in the benefit schedule and any pricing or reimbursement conditions.	Used to make decisions on inclusion in the benefit schedule and any pricing or reimbursement conditions.

			plays a role in establishing clinical guidelines.			used to negotiate pricing with manufacturer.		
<i>Key decision criteria used to establish value</i>	Size of therapeutic effect, appropriateness, and cost-effectiveness (although mainly indirectly through international and therapeutic benchmarking), burden/prevalence of disease, innovative characteristics.	Size of therapeutic effect, relevant clinical endpoints, cost-effectiveness (although unclear if and when it is considered).	Size of therapeutic effect, relevant clinical endpoints, clinical uncertainty, cost-effectiveness, quality of clinical and economic modelling evidence, budget impact, equity, innovative characteristics, ethical/legal/social considerations.	Size of therapeutic effect, clinical uncertainty, cost-effectiveness (although unclear if and when it is considered), public health impact, innovative characteristics, budget impact.	Size of therapeutic effect, quality of clinical evidence, availability of treatment alternatives, budget impact.	Size of therapeutic effect, innovative characteristics, severity of disease, availability of treatment alternatives. Cost-effectiveness/cost-benefit and budget impact generally only considered in pricing decisions.	Size of therapeutic effect, cost-effectiveness, innovative characteristics, budget impact, availability of treatment alternatives, ethical/legal considerations.	Size of therapeutic effect, cost-effectiveness, burden/prevalence of disease, severity of disease, availability of treatment alternatives, equity.
<i>Decision threshold used</i>	No	No	Yes, generally set around £20,000-£30,000 (\$31,000-\$47,000)	No (when CUA employed, the efficiency frontier is used to assess dominance)	No	No	Yes, generally set around €20,000 (\$27,000)	Yes, generally set around 500,000 SEK (\$74,000)

a This entails reviewing every new drug dossier submitted by manufacturers to support a coverage and reimbursement decisions. Therefore, in principal, manufacturers ultimately determine which drugs are reviewed.

b CEA= cost-effectiveness analysis; CMA=cost-minimization analysis; CUA = cost-utility analysis. CEA is the most widely used assessment approach, of which CUA is a type of CEA. CUA uses quality-adjusted life years (QALYs) as the principal measure of health benefit in economic evaluation, which allows comparison of the value of money of different drugs across different therapeutic areas.

c Subgroup analysis is used to explore how cost-effectiveness varies by characteristics of different patients or patient groups eligible for treatment.

Sources: Sorenson et al. (2008a); Sorenson (2010); author's compilation based on agency websites; ISPOR Pharmacoeconomic Guidelines Around the World (www.ispor.org/PEguidelines/index.asp).

Conclusions and insights for the United States

The analysis highlights that countries in Europe are becoming increasingly sophisticated in their approach to pricing and coverage policy. Governments are applying several measures in parallel to ensure national policy objectives are most effectively attained within the constraints of their respective systems. In particular, European countries are using technology assessments to inform coverage and pricing decisions, in order to ensure value for money and improve quality of care. This approach allows policy makers to focus financial investments on those drugs that offer the most value to patients and the health system, as opposed to simply those that are cheapest (or cost-neutral). Table 11 presents the different measures countries have adopted to enhance the use of evidence in drug coverage and pricing, particularly within the technology assessment process.

Table 11: *Measures to enhance evidence-based pharmaceutical coverage and pricing decisions*

Aim	Measures
<i>Ensure sufficient evidence for assessments</i>	<ul style="list-style-type: none"> • Applies risk-sharing and/or coverage with evidence development schemes (EN, FR, IT, NL) • Offers ‘access schemes’ prior to coverage review (EN) • Collaborates with manufacturers to design studies and discuss data requirement prior to coverage review (EN, IT)
<i>Strengthen link between evidence of value and coverage and pricing decisions</i>	<ul style="list-style-type: none"> • Use of a threshold to determine value for money (EN, DE, NL, SE) • Use of value-based pricing to incorporate clinical and cost-effectiveness evidence in pricing decisions and making pricing and coverage decisions concurrently (SE) • Employ evidence of therapeutic value to negotiate price of drugs (DE, FR)
<i>Monitor value of drugs post-market and disinvest of low-quality, ineffective therapies</i>	<ul style="list-style-type: none"> • Engages in post-market reviews/re-evaluation to identify areas for disinvestment or modify pricing and coverage status once a drug is used in practice (CH, DK, EN, NL, SE)

Source: Authors’ compilation

Denmark (DK), England (EN), France (FR), Germany (DE), Italy (IT), the Netherlands (NL), Sweden (SE), Switzerland (CH)

To be sure, there are challenges and outstanding problems to the pharmaceutical policies used in Europe, many of which were discussed previously.

External reference pricing, for example, is often technically challenging, in terms of selecting which countries to reference and determining the real prices of drugs. Sufficient evidence also may not be available or conclusive at the time of assessments. While new measures have been introduced in recent years to improve evidence generation, they bring other challenges. Preliminary studies on the English NHS risk-sharing schemes for multiple sclerosis and cancer drugs suggest significant challenges regarding their governance, ethics, and data collection (Briggs et al. 2010; Raftery 2010; Williamson 2010). There are also existing methodological issues with current assessment methods and concerns that they may not adequately capture all aspects of a drug's value (Drummond et al. 2009; Kennedy 2009; Oliver and Sorenson 2009), and post-market reviews of existing drugs, while important, require additional resources and are politically challenging if coverage status is modified.

However, the available evidence suggests that overall direct price control measures have been effective in containing costs (Espin and Rovira 2007; Giuliani et al. 2003; Lee et al. 2012). In addition, Cohen et al. (2013) found that relative to the approach used in the US, and in the Medicare programme in particular, the European evidence-based approach to coverage and reimbursement appears to result in reduced prices for those drugs included in national formularies (Cohen et al. 2013). The result is improved affordability for payers and increased access for patients. Cost-containment is not a main health policy objective in itself and should not be confused with efficiency, but is one of the tools that governments use to manage pharmaceuticals (Mossialos et al. 2006). To the extent that evidence of value is used to inform coverage and reimbursement decisions, however, greater efficiency may be attained.

The experiences of European countries offer several insights for the United States. Enhancing health system efficiency and quality of care as well as reducing costs are central goals of recent and ongoing health reforms. Considering the entrenched resistance to pharmaceutical price controls in the US, there are likely few lessons to be learned from international pricing systems. However, important insights can be gained from European evidence-based drug assessment approaches to support coverage and reimbursement determinations. While there is no one specific national model that is ideal or can be wholly transferred to the US, there are particular

strategies that could be modified to best meet US policy needs and fit the unique features and complexities of the American context.

In the US, there is a lack of publicly available, accessible, and robust comparative information on the effectiveness of drugs and other health interventions (IOM 2007). This gap makes it difficult for clinicians, other decision makers, and patients to make informed choices on which interventions work best and under what circumstances. CER (or, more broadly, HTA) is one viable tool to support evidence-based decision making on new drugs and to meet public expectations of safety, effectiveness, and value for money. Recent investment in PCORI offers an important step forward to better support a CER infrastructure in the US. To date, PCORI has approved 197 CER research awards totalling more than \$273.5 million.

Currently, there are no blanket prohibitions in the ACA regarding the use of CER by public and private payers⁷¹, but it remains to be seen how such research will be used. Based on the experience of Europe, the uptake and impact of CER may be limited if it does not have the authority to formally link research with policy and practice. Establishing a more formal link would improve the transparency of coverage decisions and ensure that such policies are based on independent, scientific assessment. Therefore, evidence generated from CER studies should be used to inform coverage and/or reimbursement decisions for drugs (and other health interventions). In particular, comparative effectiveness evidence could be utilized to ascertain 1) breadth of coverage (*what* to pay for), 2) depth of coverage (*how much* to pay), and 3) access rights (*for whom*). This could apply in the context of Medicare as well as in the private sector, although this is already being done in the latter to some extent.

The objective of this approach would be to use CER as a tool to move towards value-based reimbursement or, more broadly, value-based insurance. In the first instance, CER would be employed to link positive coverage decisions with reimbursement levels, where the evidence for or against a drug's comparative clinical effectiveness would be considered. Pearson and Bach (2010) proposed a

⁷¹ In the case of public payers, CER can be used for setting reimbursement rates as long as it is part of a large, "iterative and transparent process, which includes public comment and considers the effect on subpopulations" (Pearson and Bach 2010).

framework for linking CER to Medicare reimbursement, where different levels of clinical effectiveness, based on available evidence, would result in different payment levels. Drugs (and other services) offering greater health benefits than an existing alternative would receive cost-based reimbursement, while those offering only comparable benefits to an existing alternative would receive a “reference price” equal to the reimbursement rate for that alternative. New treatments for which CER evidence is inconclusive would receive cost-based reimbursement for a limited time, while further evidence is collected. If new evidence failed to demonstrate superior benefit, payment levels would drop down to the reference price. Although not commonly used by US private payers, reference prices are not foreign to Medicare; they are already used under the guise of the “Least Costly Alternative”, which could help facilitate the implementation of the aforementioned framework. Evidence on the European experience with internal reference pricing suggests that they have resulted in decreased drug prices and increases in utilization of targeted medicines, while also reducing payer and patient expenditures (Lee et al. 2012).

In order to reap the greatest potential gain from CER and strengthen its role in a broader strategy towards value-based health care, however, comparative evidence could also be applied in the above framework to define for whom a new drug works best, where it is placed in the treatment pathway (i.e. 1st line, 2nd line), and co-payment levels. For example, different payment categories could be assigned when CER evidence differed across patient subgroups. Moreover, the lowest co-payments could be used for those drugs that have the best clinical outcomes, in general or for a particular subset of patients. Similarly, those treatments with outcomes of less or highly uncertain clinical value would be granted higher co-payments. Such an approach is already being rolled out in several US health plans, with some initial evidence of resulting health gain and cost savings (Chernew et al. 2007; Choudhry et al. 2010; Spaulding et al. 2009). Public acceptance of this approach would likely be facilitated by the fact that most insurance policies have utilised some type of tiered co-payment structure for pharmaceutical benefits for decades (Wallack et al. 2004).

Of course, there are both technical and political hurdles to implementing the aforementioned framework, particularly at the Federal level. Such challenges include clearly determining evidence requirements to ascertain if a new treatment is

clinically superior; garnering resources for additional data collection and re-evaluation; modifying existing payment mechanisms to formally incorporate use of CER/reference prices; and, adding complexity to an already complex health care system. Learning from Europe, a necessary first step to address some of these issues would be arrive at some consensus on evidence standards and what level of certainty is sufficient to make judgments on superiority or ‘added value’. To allay some of the political constraints associated with Medicare, private health plans and state Medicaid programs, both of which have greater flexibility, should be encouraged and supported to be early adopters of this kind of reimbursement approach. Moreover, introduction of such measures could be done incrementally, first for only a few classes of drugs, before scaling up to more classes.

In closing, policy makers in Europe have increasingly turned to more evidence-based approaches to pharmaceutical policy. There are indeed clear differences between the US and European health systems, due in part to divergent political and historical traditions, incomes, and cultural attitudes, but adoption of some of strategies to drug coverage, reimbursement, and pricing used in Europe could potentially lend to more equitable and affordable coverage for citizens and enhance the efficiency of the US health care system. Many of these strategies are underway or under discussion in the US. Examining the European experience provides an opportunity to enhance their development, implementation, and sustainability.

Study 4: Evolving reimbursement and pricing policies for devices in Europe and the United States and considerations of value

Introduction

Given rising costs of health care and limited budgets, jurisdictions worldwide are increasingly concerned with getting better value from health care investments. This quest for value is especially evident in the case of health technologies, such as pharmaceuticals and medical devices, which account for a growing proportion of health care expenditures in almost all countries in the OECD.

The US spends more on health technology per capita than does Europe, without evidence of commensurate gains in health outcomes (OECD 2011). A range of factors influence the higher US spending, including higher prices paid for technologies; a larger volume of certain procedures, such as hip and knee implants; a greater supply or use of hospitals and doctors; and, the possibility of more readily accessible technology (Kaiser Family Foundation 2012; Squires 2012). For example, in 2009 the rate of knee replacement in the US was about 75 percent higher than the OECD median (OECD 2011).

One way Europe has been able to maintain lower spending levels is through the use of national coverage, reimbursement, and pricing policies that place greater emphasis on cost containment, efficiency, and affordability, compared to the US (Mossialos et al. 2004; Mossialos et al. 2006). Historically, European countries have employed various approaches, such as reference pricing, price volume discounts, price cuts, and centralised purchasing, to meet these ends.

Over the past ten years, however, European decision makers have shifted their focus from simple cost control to obtaining better value from investments made in new interventions (Sorenson et al. 2008a). Consequently, most European countries have established some system of technology assessment (HTA) to apply in conjunction with other policy tools. These value-based programs evaluate and weigh the available evidence on the clinical and cost-effectiveness of select interventions to

determine their value for money. The evidence is then used to inform or guide national and regional coverage and reimbursement – and, in some cases, pricing – decisions.

Although the US has traditionally failed to exercise a similar cost-conscious approach toward health technology, US policy makers and other stakeholders are increasingly focused on transforming the health system into one that seeks value, especially in light of current economic difficulties. For instance, the ACA made substantial investments in CER and in other reforms that promote value in Medicare payment and delivery systems. To date, however, comparative policy analyses and the overall discourse about health technology reimbursement and pricing in the US and elsewhere have focused on pharmaceuticals. Given the growing number and complexity of medical devices on the market, the time is ripe to examine reimbursement and pricing policies relating to those technologies. This article compares such policies in Europe and the US, with a particular focus on considerations of value. We also explore various policy initiatives, some of which have already been implemented in the US and European countries, to better support value-based device reimbursement and pricing.

Device reimbursement and pricing

Europe

In Europe, coverage and reimbursement of devices typically occurs through publicly financed national health care systems. Such systems cover approximately four-fifths of the populations of the four largest device markets: Germany, France, the UK, and Italy (Cappellaro et al. 2011). In principle, all member states are equal. Market approval of a device in one country should provide access to other markets through the CE marking process, which denotes that the device is safe and functions according to the intended purpose described by the manufacturer.

In practice, however, institutional arrangements for financing differ among countries, which can result in divergent coverage, reimbursement, and pricing decisions for a particular device (Schreyogg et al. 2009; Torbica and Cappellaro

2010). In France, for example, a centralised body makes reimbursement decisions after assessing the safety and effectiveness of individual devices. Similar bodies in England and Germany conduct broader assessments of device types or procedures and include other considerations, such as cost and cost-effectiveness. In contrast, coverage and reimbursement decisions in Italy and Spain are delegated to the various regions, which apply their own methods and requirements.

Prior to making coverage decisions, European jurisdictions typically require that high-risk, innovative, or costly devices, such as implantable technologies, undergo a HTA. An example of this process is the assessment of coronary stents by NICE in the UK. In its appraisal, the institute considered clinical trial evidence and cost-effectiveness data submitted by several manufacturers and an independent assessment group. Based on the evidence, the institute recommended use of the device only in a subset of patients at high risk of restenosis (Box 2).

Box 2: Examples of a device HTA in the UK

Background and rationale of the assessment: In October 2003, the National Institute for Health and Care Excellence (NICE) in the United Kingdom conducted a technology appraisal of coronary stents. A critical feature of the appraisal was to determine the criteria for using the newer, more expensive, drug-eluting stents (DES), as an alternative to bare metal stents (BMS).

Evidence considered in the assessment: The Institute reviewed 12 clinical trials comparing BMS with DES. Of these, seven involved paclitaxel, four sirolimus, one everolimus and one actinomycin drug-eluting stents. The evidence for the different DES was considered separately. In addition, evidence on cost-effectiveness was supplied in submissions made by the four manufacturers and an independent assessment group appointed by the Institute. The available evidence was considered by the Technology Appraisal Committee, an expert committee consisting of epidemiologists, economists, clinical experts and health service managers.

Appraisal determination: The Committee considered that, for single-vessel disease, restenosis rates were in general low using a BMS in the majority of patients requiring a percutaneous coronary intervention and therefore the routine use of a DES was not justified. However, this was not the case for patients presenting with either small-calibre arteries (<3 mm) or long lesions (>15 mm). In these patients, the risk of restenosis using a BMS was considerably higher, and that the absolute reduction in restenosis rates would justify the use of a DES. This conclusion was reinforced by the cost-effectiveness evidence. The incremental cost-effectiveness of DES, expressed as the incremental cost per quality-adjusted life-year (QALY) gained, was £94,000 for the total population of patients with single vessel disease, but below NICE's threshold of £20,000 per QALY gained for the patients with long lesions and narrow vessels.

Re-evaluation: Common to all NICE guidance, the guidance on the use of DES was reviewed five years later in July 2008. Although the clinical evidence on DES was more extensive, the major change in the intervening period was that the cost difference between the two types of stents had increased, primarily because of a greater reduction in the price of BMS. Therefore, in the revised guidance the Institute recommended that DES be used for the same sub-groups of the patient population, but only if the price difference between DES and BMS did not exceed £300.

Sources: National Institute for Health and Clinical Excellence (2003) Technology Appraisal 71. Guidance on the use of coronary stents. London: NICE; National Institute for health and Clinical Excellence. (2008) Technology Appraisal 152. Drug-eluting stents for the treatment of coronary artery disease. London: NICE.

The HTA processes required by each country differ with respect to the methods, evidence, processes, and criteria used to determine coverage (Sorenson et al. 2008a). However, all countries require that a device demonstrate therapeutic benefit, such as improved morbidity, mortality, or quality of life. In some countries, such as England and the Netherlands, evidence of cost-effectiveness is also required and measured against a value for money threshold. In some cases, the available evidence for a particular device is insufficient or inconclusive to support a coverage determination. Consequently, some European countries, including England, France, and the Netherlands, have established policies that offer restricted coverage for patients enrolled in studies designed to collect better data on safety and effectiveness. Once enough evidence is generated, the coverage decision is revisited to determine whether coverage should be extended to a broader patient population, restricted to certain patients, or removed altogether. For instance, NICE has applied its “only in

research” policy to laparoscopic surgery for colorectal cancer and endovascular stent insertion for intracranial atherosclerotic disease (Chalkidou et al. 2007).

Once coverage is determined, most European countries use prospective payment systems to determine reimbursement rates. In some cases, these payments reflect value, such as when a new diagnosis-related group (DRG) or payment amount is calculated for a new device that is based on evidence or guidance from HTA or other sources. However, because payment systems in many countries are updated infrequently, they may not adequately reimburse new technologies, especially those that are particularly innovative or costly. The lack of sufficient payment may provide a disincentive for hospitals to adopt and use new devices that may be beneficial, because the payment amount is below actual costs (Scheller-Kreinsen et al. 2011).

To address this issue, Germany, the UK, France, Italy, Spain, and Sweden have introduced separate or supplementary payments to provide partial or total reimbursements for potentially beneficial devices until they are fully captured by the payment system, either through a new DRG or an increase in the reimbursement price (Scheller-Kreinsen et al. 2011). Such payments are negotiated nationally or locally with manufacturers, hospitals, or other local authorities, and they are generally temporary, lasting two to three years. Most of the countries using this approach – particularly Germany, the UK, and France – consider evidence of therapeutic benefit and, in some cases, cost-effectiveness to determine whether a technology is eligible for the short-term payment.

Although hospitals are encouraged to collect evidence on the health outcomes and costs associated with a new medical device during the temporary payment period, there is limited evidence available to substantiate whether this is achieved. To date, these payments have been applied to drug-eluting stents, gastric bands (for weight loss), cochlear implants, and hip and knee prostheses, among other technologies. For low-risk and typically low-cost devices, such as crutches and incontinence pads, coverage and reimbursement are generally determined at the hospital level or through centralized public purchasing arrangements. In many countries, including France, Germany, and England, hospitals are increasingly entering into collaborative purchasing partnerships to negotiate lower prices, and

they are encouraged to do so by their respective departments or ministries of health. Either way, reimbursement prices are derived through reference pricing for similar existing devices or through a competitive public-tender process (Scheller-Kreinsen et al. 2011). Unlike more complex, higher-cost devices, these types of devices do not normally undergo a health technology assessment to determine value.

United States

Similar to other areas of health care in the US, coverage and reimbursement for devices are the responsibility of both public and private payers. The CMS, the largest public payer, provides coverage for the vast majority of devices once they earn approval from the FDA. After approval, most devices do not require a formal coverage determination, partly because of Medicare's prospective payment system – DRGs for inpatient care and ambulatory service categories for outpatient care. These payment mechanisms, which bundle items into an episode of care, allow payment for new technologies that offer incremental improvements over existing technologies or services. In effect, the provider simply determines coverage within the constraints of the fixed prospective payment. For a limited number of devices each year, however, CMS conducts a national coverage determination (NCD). Although there is no coherent framework for activating NCDs, this process is typically prompted by new technologies with major clinical or economic impacts – such as implantable cardioverter defibrillators – and important new evidence, substantial variation in local coverage decisions, or concerns about inappropriate use. All other explicit coverage decisions are made locally by the private insurance carriers with which CMS contracts to administer Medicare coverage.

NCDs are made through an evidence-based process, which besides CMS's own research is supported by evidence from manufacturers, physicians, and other entities, such as the Agency for Healthcare Research and Quality (AHRQ). In some cases, coverage determinations may also be made via consultation with the Medicare Evidence Development and Coverage Advisory Committee (MedCAC), which provides independent and expert advice to CMS on various clinical issues. Such evidence is used to determine the degree of benefit conferred by the devices compared to standard treatment alternatives. Unlike processes in some European

countries, NCDs do not explicitly require or consider evidence of cost-effectiveness, which has proved politically controversial (Chambers et al. 2010).

Despite these procedures, designed to improve CMS's ability to make informed decisions about the underlying value of a technology, existing evidence suggests that in the majority of cases, positive NCDs are based on poor or limited evidence from clinical studies (Neumann et al. 2008). However, in cases where technologies offer promise, but have been inadequately studied to support a NCD, the CMS can approve coverage of a device under a clinical trial or another protocol, such as an observational study or patient registry, until the required evidence is amassed. This approach, called "coverage with evidence development", has been applied to a few devices to date, including implantable cardioverter defibrillators, angioplasty of the carotid artery with stenting, cochlear implants, and left ventricular assist devices.

For local coverage determinations, depending on the technology or service considered, local administrative contractors make decisions relying on an evidence base that ranges from no evidence to peer-reviewed RCTs (Foote et al. 2004), which may be one reason why local contractor coverage often varies considerably.

The amount paid by Medicare is determined through a prospective payment for an episode of care or a retrospective payment for an episode of care or a retrospective fee-for-service payment for the actual service, or device, provided. With limited exceptions, CMS does not currently consider a device's comparative effectiveness or its cost relative to alternative treatment options in its pricing. Rather, payments are based on estimates of average cost for the provision of the particular device or bundle of care.

Similar to the situation in Europe, some beneficial yet costly new devices used in patient care may be granted separate "add-on payments" to account for the high cost of new technology relative to the base DRG payment and to encourage providers to adopt the technology (Clyde et al. 2008). To receive these payments, devices must be new and high cost, and they must substantially improve the diagnosis or treatment of beneficiaries, compared to existing treatment alternatives. In such cases, the

devices might offer a treatment option for patients unresponsive to current therapy, diagnose conditions that are currently undetectable, provide meaningful impacts to patient management, or substantially improve clinical outcomes. Similar payments, called “pass-through payments” and employing the same eligibility criteria, are used for devices provided in the outpatient setting. For add-on payments, Medicare pays an amount equal to 50 percent of the additional costs of treating a case using the new device, which is capped at 50 percent of the estimated cost of the new technology. Pass-through payments are made equal to 100 percent of the reported costs of the new device minus the device costs already built into the base payment rate. A device is eligible for an add-on or pass-through payment until data reflecting its costs are used to recalibrate the appropriate diagnosis-related group weights, generally two to three years after the new technology has entered the market.

Add-on payments have been extended less frequently than pass-through payments. Fewer than ten technologies have been approved for add-on payments, while pass-through payments have been made for more than a hundred different device categories. The majority of add-on payments made to date have been for implantable medical devices, while a wider range of devices have received pass-through payments.

Private payers cover about two-thirds of the US population. There is considerable diversity among insurance plans’ coverage and reimbursement policies. Private payers sometimes look to CMS NCDs to guide their decisions, but largely develop independent policies based on the goals of the individual plan. Private insurers also tend to make coverage decisions more quickly after FDA approval than does Medicare, although the speed of decisions made by private payers depends on the amount and quality of evidence of clinical benefit (Basu and Hassenplug 2012).

Private payers are increasingly considering evidence of value to support formulary and tier placement decisions and in applying preauthorisation or utilisation reviews. For example, WellPoint draws on comparative effectiveness evidence and on input from panels of medical experts to assign existing and new treatments to one of four value tiers (Academy Health 2010). Both the Blue Cross Blue Shield Association and Kaiser Permanente have established institutional policies and

dedicated funding for in-house or external programs that generate evidence to support coverage determinations and clinical practice guidelines. Other, smaller health insurers and health plans often rely on independent research organisations to provide evidence reports on new devices and other technologies. Similar to Medicare, when evidence is considered, private payers tend to consider effectiveness, not costs or cost-effectiveness (at least not explicitly).

Private insurers rarely directly reimburse devices. Rather, insurers negotiate payment terms directly with physicians and hospitals, where each medical procedure or episode of care is reimbursed at a specified or negotiated amount that must cover the price of the device along with other items – such as supplies, labour, and facility costs – that are part of the procedure or care episodes. Negotiated reimbursement amounts are rarely based on whether a technology is more effective, or easier or more efficient to use, than existing treatment alternatives.

United States versus Europe

As highlighted in the overview, there are distinct differences in the approaches that Europe and the US take toward reimbursing and pricing medical devices. European countries have more centralised processes for making coverage determinations than the US, which has a patchwork of public and private payers that may employ different processes and criteria to make decisions.

Moreover, compared to the US, Europe more formally and consistently considers value to determine which technologies to cover, especially complex, costly ones. In the US, a limited number of devices actually undergo a formal value assessment at the time of a coverage decision, especially within the public sector. Europe also places more emphasis on accounting for cost-effectiveness. In the US, cost-effectiveness raises concerns about the formal rationing of care and whether such analyses can adequately capture the value of interventions for different population subgroups.

There are similarities, however. Both the US and Europe tend to use evidence of value more frequently to support coverage decisions than to guide reimbursement or price decisions. However, the US and many countries in Europe have introduced temporary payment mechanisms to provide increased reimbursement for beneficial, but costly, technologies. These approaches aim to allow payers to balance the goals of ensuring adequate payment for beneficial new technologies and being prudent purchasers. For the selected number of technologies that receive such payments in the US and Europe, evidence of therapeutic benefits play a central role in determining eligibility. Costs are also considered, with a number of European countries also accounting for cost-effectiveness. Finally, where evidence is applied in coverage policies, both jurisdictions are often faced with having limited information to inform decisions.

The lack of high-quality evidence for making informed coverage decisions means that coverage may be provided for a new device based on fair or poor evidence or that access to potentially beneficial technologies may be delayed or denied until better evidence is available. Conditional coverage with evidence generation has therefore gained some use in recent years.

Policies to improve value-based device reimbursement and pricing

We outline and discuss potential initiatives to obtain better value in health care in Europe and the US, highlighting their possible advantages and disadvantages.

Fostering pre- and post-market evidence

One of the main challenges in ensuring adequate evidence of effectiveness to make coverage and reimbursement decisions is that such data are not generally required for market approval. Following recalls of articular surface hip prostheses and PIP breast implants, however, the FDA and European regulators are now considering an overhaul of the current regulatory frameworks for medical devices with a particular focus on strengthening premarket requirements for high-risk technologies (European Commission 2012; Institute of Medicine 2011).

One action that both European and US regulators should consider is raising the pre-market evidence requirements for new devices. Current requirements allow clinical evaluations of most new devices to be based on similar existing (predicate) technologies rather than the actual device in question, and the clinical data submitted to be based on a literature review along. Current systems therefore reward ‘fast followers’ that can take advantage of existing evidence about similar products that are already on the market.

Instead of simply assuming that devices of a given type are equivalent, fast followers could be required to generate the same level of evidence as exists for other devices already on the market. Discussions could take place between regulators and the first manufacturer to determine the level of evidence required—for example, a registry or a randomised controlled trial. Imposing such a requirement would not only give industry an incentive to undertake clinical studies on new devices and foster a better understanding of the comparative differences between devices, but it would also enhance public health protection.

Regulators should supplement efforts to strengthen premarket evidence with incentives and, where possible, requirements for post-market evidence generation. Pre-market evidence is often not ideal for ‘real-world’ decision making because of uncertainty regarding long-term outcomes, effectiveness in different practice settings, and benefits and risks to populations that are not well represented in clinical trials.

The various approaches to coverage with evidence development for new technologies in Europe and the US offer some opportunities to ensure that sufficient post-market evidence is available to inform coverage determinations. Although CED has been used on a limited basis, it has provided evidence that otherwise might not have been obtained. However, substantial improvements to this approach are needed. Because it has been used on a limited basis, clear and predictable criteria for its application and methods are lacking. There are also challenges in delineating well-defined funding sources to cover the large research costs and an infrastructure to collect and share data.

CED should be aligned with existing mechanisms to expand electronically available health data, including longitudinal patient registries; electronic health records; and, in the United States, claims data collection and analyses. Some European countries— including Germany, Italy, Sweden, and the UK —have introduced registries, particularly in orthopedics and cardiology, to collect post-market data. Typically, medical associations, academic centers, and national research organisations collaboratively support these registries.

Similarly, the PCORI and the NIH in the US could provide support for an ongoing infrastructure for registries or clinical trials in major clinical areas. The involvement of clinicians or medical associations may prove particularly helpful, given their early contribution to device development and acquired early knowledge of particular technologies (Wilmschurst 2011). These efforts would help ensure that the necessary data are generated to support CED schemes and, ideally, that better evidence exists to make informed coverage decisions in the first place.

Given the substantial expense and time involved in collecting reliable data on new technologies, more public-private collaboration would be desirable. One approach would be for payers and regulators to provide scientific advice and manufacturers to ensure that clinical studies meet the evidence requirements for both market authorisation and coverage and reimbursement. In relation, concurrent review of devices by regulators and payers could help reduce evidence generation burdens and thereby allow beneficial technologies to reach patients more quickly. In the US, for example, the FDA and CMS have initiated a voluntary, two-year “parallel review” program for devices, which entails a partial alignment of their respective review processes for regulatory approval and coverage, respectively (Messner and Tunis 2012).

Exploring new approaches for assessing value

Another possible initiative would be to establish new methods for assessing the value of devices. Devices have particular characteristics that introduce unique challenges to measuring their value (Drummond et al. 2009; Sorenson et al. 2011a). For instance, devices undergo frequent modifications following initial development,

which means that they do not “stand still” long enough during the period of randomised controlled trials to allow for adequate data collection. Moreover, accurate or effective use of devices often depends on the skills and training of the health professionals who use them, especially for those devices used in surgery (Sorenson et al. 2011a). Practitioners may acquire more expertise with a device over time—or move along the “learning curve”—even over the period of a trial (Guillou et al. 2005).

Some of these issues can be tackled through the use of tracker trials, which begin in the early stages of technology development and follow the evolution of a device (Lilford et al. 2000), but these are not commonly conducted. Although not officially required by the regulator, the trial of endovascular aneurysm repair, a procedure using a stent, is an example of a trial using this approach (Brown et al. 2012).

Alternative study methods might also be better suited to medical devices. Although randomised controlled trials are considered the gold standard, there is increasing recognition that alternative study approaches may be suitable in some instances. For example, the CER initiative in the US has focused attention on pragmatic randomised controlled trials, which take place in real-world practices as well as observational studies and patient registries (Dreyer et al. 2010; Chalkidou et al. 2012).

Linking evidence of value to reimbursement

Value-based reimbursement, an approach increasingly of interest to US private payers, may provide a viable option to better incorporate evidence into reimbursement decisions. In a survey of employer-sponsored health plans, Niteesh Choudhry and co-authors (2010) estimated that 81 percent of large employers plan to offer this approach in the near future.

This approach sets different reimbursement rates for different levels of clinical effectiveness, based on available evidence. It may also entail differential copayments for treatments of demonstrated high value versus those of questionable or low value,

and it aims to encourage the use of services when the clinical benefits exceed the cost. The approach may likewise discourage use when the benefits do not justify the expenditure. This may also help remove financial barriers to beneficial technologies and thereby increase patients' compliance with treatment (Maciejewski et al. 2010), which in turn can improve health outcomes, rein in costs, and assist in controlling total spending by health plans or hospitals. Elements of this approach could also be employed in Europe, but given the absence of cost sharing, evidence of value could not be tied to copayments.

A related strategy that could be considered is the use of performance-based reimbursement and pricing strategies that link payments to patient outcomes. For example, a certain reimbursement price may be set—and later modified—according to whether the device is used in accordance with evidence-based clinical guidelines or produces satisfactory clinical outcomes. Using such strategies, payers may face less financial risk from the treatment of demographically different patient groups that were not included in clinical trial testing or that did not demonstrate substantial improvement (Towse and Garrison 2010). This approach has been used on a limited basis in Europe and by private payers in the US, but only with regard to pharmaceuticals, not devices.

These approaches need to be applied with care, however. The few performance-based schemes implemented for pharmaceuticals in Europe have been costly to administer and marked by difficulties regarding oversight, methodological requirements, and ethical considerations (Raftery 2009). Such challenges may be more pronounced in the case of devices. In addition, even in cases where the available evidence demonstrates that a device provides low value, it may prove administratively and politically difficult for payers to disinvest from the technology once it has diffused into practice (Elshaug et al. 2007).

Conclusion

Policy makers and other stakeholders in Europe and the US are increasingly concerned with getting value from investments made in technological innovation. One potential solution is to rely more heavily on studies of the effectiveness and costs of

new technologies to inform coverage, reimbursement, and pricing decisions. Historically, such efforts have largely focused on pharmaceuticals. But with the growth in the number and complexity of devices, the US and Europe have shown interest in applying evidence of value in coverage and reimbursement decisions, albeit with varying degrees of implementation and success.

Although these strategies are still unfolding, we have outlined a number of them that could help support the timely generation of evidence to inform value-based decisions about reimbursement and pricing for devices. Further discussion and research are needed on the various options to substantiate their effectiveness, best practices, and areas for improvement.

**SECTION III: IMPLEMENTATION CHALLENGES
ASSOCIATED WITH POLICIES TO IMPROVE HEALTH
TECHNOLOGY REGULATION**

Study 5: The evolution and impact of health technology assessment in Europe

Introduction

Few issues in health policy have generated more commentary, and perhaps more controversy, than HTA. Over the last 30 years, several European countries have established agencies and programmes to carry out HTA; there is now a flourishing research sector to support assessments; and its use in decision making has provoked much discussion within academic and policy circles. Although HTA was initially considered to be primarily an academic exercise, it is now used to support coverage and pricing decisions; inform the development of clinical practice guidelines and quality standards; develop public health programmes; and aid purchasing or disinvestment decisions (Sorenson et al. 2008a).

While HTA has advanced most rapidly within Europe, it had its origins in the US with the OTA in the early 1970s, which sought to provide policy makers with information on the economic, social and legal impacts of modern technology (Goodman 2004). The OTA was later disbanded principally due to political reasons, but it served as a model for other countries, namely those in the European community (OTA 1996). Throughout the 1980s and 1990s, the soaring costs of health care and emerging concerns about the effectiveness of many existing medical practices and variation in access to care in Europe spurred interest in HTA (Oliver et al. 2004). With governments (and other decision makers) increasingly required to efficiently and equitably allocate resources among available health-care technologies, it was recognised their costs and benefits needed to be assessed. HTA was also considered a potential mechanism to help policy makers justify or legitimise rationing decisions and render the process more transparent. The first institutions or bodies dedicated to HTA were established in France and Spain in the early 1980s and in Sweden a few years later in 1987 (Velasco-Garrido and Busse 2005). These early efforts aimed to provide unbiased, scientifically rigorous assessments of health technologies and other interventions (e.g. procedures) for a wide range of end users, including policy makers, payers, physicians and patients. A decade or so later,

several other countries, including England and Germany, instituted new agencies or programmes to carry out HTA. In recent years, several Central and Eastern European countries, such as Poland and Hungary, have followed suit. HTA programmes and activities have also been established at regional and local levels in many countries (Velasco-Garrido et al. 2008).

HTA in Europe has undergone notable growth and change since the initial establishment of these institutions and programmes. This paper reflects upon this evolution and discusses key developments over the last 10 years, with a focus on England, France, Germany and Sweden. All of these countries have significant experience or a long history of using HTA and have influenced the approach used in other countries. In particular, we discuss how HTA has evolved over time in these countries, in terms of organisation and governance, objectives and scope, processes and procedures, stakeholder involvement, assessment methods, applications to decision making and implementation; its impact on various outcomes; and key similarities and differences between countries. We focus principally on the following national bodies: NICE (England); the Institute for Quality and Efficiency in Health Care, IQWiG (Germany); French National Health Authority, HAS (France); and the Dental and Pharmaceutical Benefits Agency, TLV (Sweden).⁷² Existing comparative research highlights that while these national agencies share some of the same basic objectives, there are divergences in how HTA is configured and operates within the health-care system (Chalkidou et al. 2009; Sorenson et al. 2008a; Velasco-Garrido and Busse 2005).

Key developments in HTA

Organisation and governance

The organisation and governance of HTA in the various countries has witnessed significant changes over the last 10 years or so. In general, there has been an expansion in the number of HTA bodies and breadth of activities in each country,

⁷² See Velasco-Garrido et al. (2008) for a comprehensive overview of other national HTA organisations and activities.

with one or more entities assuming a key regulatory or advisory role⁷³ in policy making, typically with regards to coverage and reimbursement and, sometimes, pricing decisions. In some countries, such as France and England, there has been some degree of reorganisation to consolidate and rationalise assessment functions.

In Germany, HTA came to the fore principally over the last 10 years. The 2000 health-care reform law (Statutory Health Insurance Reform Act of 2000) first established a Coordinating Committee responsible for technology assessment in the hospital sector (IQWiG 2010a). The law also charged the German Institute for Medical Documentation and Information (DIMDI) with the administration of a federally funded HTA programme to provide evidence in support of decision making processes in health care (Perleth et al. 2009). The German Agency of Health Technology Assessment (DAHTA) was later created within DIMDI. Another round of reforms in 2004 (SHI Modernization Act) established the Federal Joint Committee (GBA) to serve as the supreme decision-making body of the ‘self-governed’ health-care system; its directives define the provision and reimbursement of a wide range of health technologies as well as non-medical treatment within the benefit package (Gerhardus 2006). In the same year, IQWiG was founded to serve as an independent advisory body that conducts assessments on the benefits and harms of medical interventions (IQWiG 2010a). One reason for IQWiG’s creation was to strengthen the use of evidence-based medicine and to aid the GBA’s appraisals based on ‘the principles of effectiveness, necessity and cost-effectiveness’ (Perleth et al. 2009).

Similarly, in Sweden, the Pharmaceutical Benefits Board (LFN, later the TLV after expanding its remit to include dental care) was introduced in 2002 following reforms aimed at enhancing the cost-effective use of pharmaceuticals and ensuring equal benefits across the country. The TLV significantly changed the introduction of drugs in Sweden; rather than automatic reimbursement within the benefit scheme, reimbursement and pricing decisions were based on an assessment of available evidence, with resultant appraisal underpinned by the principles outlined in the 1982 Health and Medical Service Act – health dignity, need and solidarity, and cost-effectiveness (Sorenson et al. 2008a).

⁷³ A regulatory role involves direct decision-making authority, whereas an advisory function entails making recommendations to a governmental body.

While HTA has always been highly centralised in France, in 2004 the French National Health Authority (Haute Autorité de Santé, HAS), established by the Health Insurance Reform Act, replaced the previous National Agency for Accreditation and Evaluation and assumed the directorate responsible for HTA at the French Medicine Agency, in attempts to place all activities focused on improving quality of care under the remit of one body. Besides ensuring greater efficiency, the reorganisation was driven by a desire to facilitate equity within the health-care system. HAS' responsibilities are diverse, including technology assessments, best care standards, guideline publication and health organisation and professional accreditation.

The breadth of organisations involved in HTA in England has always been diverse. During the 1970s and 1980s, a growing number of bodies, from the Medical Research Council to the Department of Health and universities, became involved in HTA. Later, in the early 1990s, a national HTA programme was established to set research priorities, commission HTA studies, and disseminate reports. During this time, there was increasing concern over national variations in care ('post-code prescribing') and a desire to 'de-politicise' decisions about which technologies to cover in the NHS. In response to these issues, NICE was established in 1999 to serve as an arms-length organisation that employs the best available methods to provide national guidance on health technologies, interventional and diagnostic procedures and clinical guidelines. In 2004, NICE assumed the responsibilities of the Health Development Agency, expanding its remit to include the evaluation of public health interventions (e.g. preventive screenings, smoking cessation). Following recent transitions in government, NICE's role in the NHS is set to expand further in the next year or so, with a role in social care and responsibility for setting quality standards (e.g. stroke, dementia), which will feed into various commissioning activities, regulation and pay-for-performance schemes. At the same time, other functions of NICE's work may well evolve, with the announcement of the move towards value-based pricing (Department of Health 2010).

There are some important differences in the role these organisations play in HTA. In England and Sweden, NICE and the TLV, respectively, are regulatory

bodies with direct decision-making authority and are involved in both the assessment and appraisal process. In contrast, IQWiG in Germany is positioned further downstream in the decision-making process, where it assumes an advisory role, as does the French HAS, an arrangement arguably influenced by their social insurance-based health-care systems. IQWiG and HAS conduct assessments and make recommendations on coverage and reimbursement, generally regarding pharmaceuticals, to the GBA and Ministry of Health, respectively, who appraise the evidence and render a final determination. HAS also advises individual insurance plans responsible for decisions regarding other types of health services. As a result of their regulatory function, the workings of NICE and the TLV are more closely linked to, and therefore influential, in the policy-making process, which we discuss further in the following sections.

Objectives and scope

The objectives guiding assessment requirements and methods have changed over time. All four agencies were established to improve the effectiveness and quality of health care through their activities, but only NICE and the TLV also focused on efficiency or value for money, based on cost-effectiveness analysis, as an explicit aim of HTA. Recent reforms in France and Germany, however, have introduced such objectives to HAS and, to some extent, IQWiG (Chalkidou et al. 2009; Chevreul and Durand-Zaleski 2009).

In 2008, the French Social Security Finance Act introduced the consideration of value for money in HAS's reviews and recommendations. A HAS Commission for Economic Evaluation and Public Health (Commission é'valuation é'conomique et de sante' publique, CEESP) was established to oversee the integration of cost-effectiveness into public decision making as well as clinical practice. The CEESP will issue a recommendation on a drug's cost-effectiveness, which will be considered alongside advice regarding reimbursement and pricing. However, cost-effectiveness will only be considered when reassessing classes of drugs, medical devices, or organisational aspects of health delivery already in use; new technologies will continue to be evaluated primarily on therapeutic benefit. To date, the use and

implementation of this new approach is uncertain and currently under discussion (Chevreul and Durand-Zaleski, 2009).

A new German law (Neuordnung des Arzneimittelmarktes, AMNOG) enacted in January 2011 now requires an early assessment of the benefit of medical products. Following regulatory approval, the GBA will assess whether the new medical product demonstrates additional benefit in comparison to a corresponding established therapy. Assessment of added benefit will be based largely on manufacturer data (e.g. clinical studies) and principally relate to relative effectiveness, although costs of the therapy will also be considered. Evidence of cost-effectiveness will only be sought if a price cannot be negotiated between the manufacturer and the sickness funds based on the GBA relative effectiveness and cost–benefit assessment. Chalkidou et al. (2009) suggest that such changes in both countries were driven by the recognition that the lack of early and comparative evaluation was limiting comprehensive assessments, and by the growing need to prioritise expenditure across different types of health technologies.

The scope of HTA activities undertaken by these agencies has also evolved. While the focus on assessments has been principally on pharmaceuticals, this has expanded over recent years to include other technologies and interventions. For example, NICE assesses not only drugs, but also public health interventions, surgical procedures, and has recently established a separate technology appraisal programme for medical devices (Sorenson et al. 2008b). HAS also evaluates a wide range of health-care interventions. Expanding HTA to non-pharmaceuticals, however, has introduced new procedural and methodological challenges (Sorenson et al. 2011a). In the case of medical devices, for example, application of existing evidence standards – developed with pharmaceuticals in mind – can be problematic, as randomised controlled trials are often unavailable or too small to detect meaningful differences in clinical and economic outcomes between technologies. Moreover, devices are often developed iteratively, so there is unlikely to be a substantial ‘steady state’ period when all relevant evidence can be collected and evaluated.

Processes and procedures

Increasingly over the last decade, it has been debated that good process in HTA is important, at least from an instrumental perspective (Culyer and Lomas 2006). While the notion of ‘best practices’ for conducting assessments continues to evolve, consensus is coalescing around several dimensions (Drummond et al. 2008; Chalkidou et al. 2009). Good process entails independence for those conducting assessments; transparency; openness about those involved in the assessment process; explicit timelines for completing assessments; opportunities for stakeholder input; and clear rules for appealing decisions.

While it is beyond the scope of this paper to evaluate each country on these principles, it is possible to note several developments in this area. As discussed earlier, most HTA bodies operate with a degree of independence from government and recognise and strive to protect against any conflict of interest in their processes and policies (Chalkidou et al. 2009). In England, NICE has increasingly made its processes and guidance publicly available to improve transparency. With the same aims in mind, IQWiG publishes previously unpublished (manufacturer) data considered in benefit assessments (IQWiG 2008) and the TLV publishes its decisions on its website. Both England and Sweden have also instituted formal appeals processes. Finally, all four agencies afford opportunities to stakeholders to contribute input on various aspects of the assessment process, which we discuss subsequently in further detail, with some countries, such as Germany, making it a legal requirement.

Although steps have been taken to improve the HTA process, there are potential trade-offs involved. For example, opportunities for stakeholder input and appeals can delay assessments and resulting decisions (Haycox 2008). Greater transparency can leave the HTA process open to challenge by stakeholders. Additionally, it can be argued that procedural aims or principles are often context-specific and therefore adhering to one particular set of expectations and practices across systems or even decisions may not be desirable or appropriate.

Stakeholder involvement

As highlighted above, countries have sought to increase stakeholder involvement in various aspects of HTA in recent years as a way to enhance the relevance, acceptance, transparency and legitimacy of the process (Abelson et al. 2007; Milewa and Barry 2005; Milewa 2006; Syrett, 2003). In addition, stakeholder involvement was seen as a way to acquire important information on new treatments and people's medical preferences and values. Today, various stakeholder groups, including industry representatives, health professionals and patients and patient groups are participants in HTAs.

Manufacturers have assumed an increasingly integral role in HTA processes. In all countries, they submit evidence to be used or considered in assessments. In England, France and Germany, industry representatives also serve on assessment or appraisal committees that review the available evidence and develop recommendations. Given the growing number and diversity of new technologies, which often add more complex evidence and assessment requirements, HTA agencies have looked to work with industry more closely to outline evidence requirements and scope assessments. For example, in 2008, IQWiG launched its programme 'IQWiG in dialogue', which aims to offer representatives from the scientific community, industry and the Institute the opportunity for scientific and technical discussions on various topics related to the work of the Institute (IQWiG 2010b). Similarly, health professionals, medical associations and payers are often involved in assessment or appraisal committees. The GBA, for example, includes members from associations of physicians, hospitals and sickness funds, along with patient representatives, as does many of IQWiG's governing boards and advisory panels (Nasser and Sawicki 2009). NICE recently established a comparable body, a Partners Council, to engage with industry, NHS managers, clinicians and academics to discuss strategic challenges (Chalkidou et al. 2009). It also offers Scientific Advice, a fee-for-service consultation to manufacturers to ensure their early product development plans generate relevant evidence for future submissions to the institute. All of these stakeholder groups, in addition to manufacturers, patient groups and the general public, are also offered opportunities to identify topics for assessment in most countries.

Historically, patients, consumer groups and the public have been the least involved or represented in the HTA process, but this has changed over time. In England, for example, NICE established a Citizens Council in 2002 to gather public perspectives on key issues that inform the development of the Institute's guidance. The council was partly a response to criticism that the ethical and social aspects of decisions were not being sufficiently articulated and that NICE was recommending too many new drugs, sometimes heavily influenced by industry and interest groups, leading to aggravated forms of implicit rationing elsewhere (Cookson et al. 2001; Devlin et al. 2003). The council is comprised of lay members reflective of the broader population; health professionals, industry and interest group representatives are precluded from participation. To date, the council has discussed such issues as clinical need, age and rare conditions. In recent years, HAS and IQWiG have established similar opportunities for these stakeholder groups. HAS, for example, holds patient and consumer focus groups and public debates to explore sensitive and controversial topics and these groups may be members of document revision groups, which aim at improving the comprehensibility of guidance.

However, stakeholder involvement is not without challenge. For example, it has been argued that the opinions of clinicians and other professional actors often take precedence over more 'lay' representatives in assessment committee or working group meetings, especially given the latter almost never argue against health technologies and more frequently make more emotion-driven arguments (Milewa and Barry 2005; Milewa 2006). This is due, in part, to the fact that patient group representatives, for example, are often personally invested in ensuring access to a particular technology or treatment, but also because they are sometimes affiliated with or funded by pharmaceutical companies⁷⁴. Milewa (2006) suggests that the perspectives of 'non-professionals' are thus more likely to have their credibility or legitimacy questioned.

In relation, a growing role for industry in the HTA process raises concerns about the potential for 'regulatory capture', where manufacturers influence regulators through lobbying, superior information (i.e. about the products under assessment),

⁷⁴ Only HAS requires industry to declare funding of patient groups.

and potentially other means to guard against unfavourable regulation. Other invested stakeholders, such as patients and the general public, however, face major obstacles, such as limited resources and a lack of information, to exercise influence (Olson 1965). Consequently, decisions may ensue against the interests of a broader range of stakeholders and biased in favour of industry, a highly resourced group with significant stakes in resulting policy (Schlozman and Tierney 1986).

A broader concern is that, overall, increased stakeholder involvement has not substantially resulted in greater consideration of a broad range of views and values, but rather serves to merely justify or legitimise difficult decisions or those that would be taken anyway. Indeed, explicit resource allocation can heighten the politicisation of decisions (Oberlander et al. 1994). Therefore, stakeholder involvement may be a mechanism to minimise potential challenges regarding their rationality and legitimacy and the assessment process in general.

Assessment methods

There have also been several key developments in assessment methods over the last 10 years. Generally, assessments can involve different methodological components, including a clinical effectiveness assessment (i.e. therapeutic benefit, safety) and an economic evaluation, typically involving cost-effectiveness analysis. They can also evaluate broader social and ethical impacts.

In England, NICE has been recognised for its methodological rigour and, in particular, its assessment guidelines, which set standards for consistency and compatibility of the studies submitted to NICE. One of the more contentious aspects of the guidelines is NICE's requirement for the use of QALYs to measure health benefit (NICE 2008a). Although the QALY is broadly accepted, there is some debate regarding its limitations, namely that there are methodological problems with its use and that it does not fully capture the social value of interventions (Johnson 2009; Oliver and Sorenson 2009). Drummond (2009) also notes that NICE has been influential on other methodological approaches, most notably in the use of mixed treatment comparisons in the use of probabilistic sensitivity analysis to account for

uncertainty surrounding cost-effectiveness estimates. The TLV in Sweden also adopted and uses the QALY approach.

In contrast, HAS uses a expected service (*service attendu*, SMR) and added value (ASR) ranking to determine the coverage and pricing, respectively, of new drugs and devices. Technologies are evaluated alone to determine their therapeutic benefit using the SMR rating and then against comparator products using the ASR ranking, where they are graded from I to V, with I indicating a major improvement over existing treatments and V denoting no therapeutic benefit or added benefit compared to existing alternatives. HAS has also developed a new ‘social benefit measure’ (*service rendue à la collectivité*) to allow for assessments to not only evaluate the therapeutic benefits of interventions, but also economic endpoints and important ethical, social and legal considerations (National Health Authority 2007). While its use to date has been limited to screening programmes, ongoing discussions suggest that it may be applied to pharmaceuticals and other interventions.

Germany has also taken a different approach. In 2009, IQWiG, in line with its expanded remit to consider costs, adopted an alternative approach, ‘efficiency frontier analysis’. This method, which has generated considerable debate within the scientific community in and outside of Germany,⁷⁵ uses prior funding decisions for similar products in the same disease area to determine the maximum ceiling reimbursement level for a drug. However, no decisions using the efficiency frontier have been made to date and IQWiG’s recommendations continue to rely almost entirely on cost-minimisation analysis, which does not assess efficiency.

In response to criticisms that the time taken to assess technologies⁷⁶ poses barriers to patient access to new treatments, HTA bodies in England, France and Germany have established new expedited processes and methods for evaluating certain categories of drugs, typically those that are deemed highly innovative or for those treating life-threatening illnesses. England, for example, is increasingly

⁷⁵ Critics argue that the efficiency frontier approach fails to reflect international standards of economic evaluation and that it potentially hinders effective resource allocation, as it does not allow decisions to be made across different therapeutic areas. However, one reason underpinning this approach is that it is not possible under German law to deny drugs above a set standard threshold; rather, the aim is to determine a fair price that reflects additional clinical benefit (Bundesgesundheitsministerium 2008).

⁷⁶ On average, assessments take between three months to two years, depending on the HTA body, drug, indication and extent of available data.

conducting more single technology appraisals (STAs), which consider only a single technology in a single indication and rely principally on manufacturer submitted data, rather than a de novo analysis. Research on the impact of STAs indicate that they reduce the time to publication of guidance by about eight months compared to the usual multiple technology appraisal process (Casson et al. 2013). However, appeals against the final appraisal determination have more than doubled the time it takes for STAs to conclude. In Germany, the new reforms will require manufacturers to submit their evidence dossiers to the G-BA and IQWiG and agree on a price within one year. However, it is important to note that greater timeliness may come at a price, in the form of less robust assessments and reduced stakeholder input.

Applying evidence to decision making

Evidence generated by assessments is subsequently used to support decisions (appraisal) on the value of new technologies. Countries differ, however, in terms of what criteria are considered in the appraisal process. All countries, for example, prioritise evidence on therapeutic benefit (Sorenson et al. 2008a), but to date, only NICE and the TLV explicitly consider cost-effectiveness. Cost-effectiveness data is appraised using a decision threshold, measured in cost per QALY, to support determinations of value for money. The threshold differs between countries and is often implicit and case dependent. It has been estimated that NICE's and TLV's threshold ranges from £20,000–£30,000 to £45,000–£50,000, respectively (Devlin and Parkin 2004; NICE 2008a; Persson and Hjelmgren 2003; Rawlins and Culyer 2004).

However, there have been recent debates around whether the threshold should be raised (or lowered) and if a broader range of decision criteria, reflecting political, social and ethical considerations, should be more formally integrated into decision making (Kennedy 2009; NICE 2008b; Raftery 2010a; Towse, 2009). Discussions around these issues stemmed partly from concerns that certain patient populations (e.g. those with rare or orphan conditions or terminal disease, elderly and disabled) were disadvantaged by existing decision rules and that the value of innovation was not accurately captured by current approaches. In recent years, both NICE and the TLV have considered adopting or have applied a revised approach to the threshold in

certain circumstances. Sweden, for example, has discussed adjusting the threshold to better account for need or equity considerations, decision criteria outlined in the Board's founding principles, especially for drugs that are potentially expensive or address unmet medical needs. NICE recently extended its threshold for drugs aimed at end-of-life care under some circumstances to facilitate cancer drug access in the NHS, albeit under intense media, ministerial and public pressure to do so. While decision makers do consider other criteria (e.g. disease burden, equity, innovation), at both the assessment and appraisal stage, it is often unclear (and lacking transparency) what aspects are indeed considered and with what weight (Kennedy 2009). This is likely in part reflective of the need and desire of decision makers to maintain a degree of flexibility to consider political influences or other context specific factors of importance.

The different national approaches to assessments and the appraisal process begs the question whether HTA bodies arrive at the same or different decisions. A recent study by Kanavos et al. (2010) examining all decisions made by six HTA agencies between 2007 and 2009, including NICE, HAS and the TLV, found a significant degree of heterogeneity across the coverage decisions made by the different agencies, with outcomes differing in more than half of the cases. A closer examination of agency decisions across similar cancer, central nervous system and orphan drugs showed a number of factors driving the differences in decisions, including divergent clinical and economic evidence requirements, preferred clinical endpoints, data interpretation, choice of comparator and use of cost-effectiveness thresholds. For example, compared to NICE, the TLV was more willing to consider need over cost, as intimated earlier, approving drugs beyond the threshold, up to £75,000, when there is high clinical need in certain sub-populations. Other research (Bending, Hutton and McGrath 2010; Patel et al. 2010) on cancer drugs substantiates the influence of evidence requirements, where agencies (i.e. NICE) focused on cost-effectiveness less frequently gave positive recommendations (about 50%) compared with those only requiring evidence of clinical effectiveness (i.e. HAS).

Data availability and quality may also play a role in differential decisions across countries, introducing uncertainty into the appraisal process and opportunities for misinterpretations of existing evidence. In response, all four HTA agencies are

experimenting with new approaches to generate better pre-and post-market data. For example, all countries have recently employed some form of coverage with CED, risk-sharing agreements, or patient access schemes, where coverage of a technology is made conditional based on arrangements for additional post-market evidence collection or meeting certain health or financial outcome targets. Concurrently, agencies are assuming greater involvement in prospective data generation (Chalkidou et al. 2009). NICE, for example, does so through its ‘only in research’ option of conditional reimbursement (Chalkidou et al. 2008), with a similar programme in use in Germany and under consideration by HAS in France. Such programmes condition the use of an intervention to those patients receiving it as part of a well-designed programme of research, and are typically applied when there is insufficient evidence to make a conclusive coverage or reimbursement decision. These approaches, however, often require more sophisticated data collection and assessment methods and additional time and resources, thereby potentially adding further complexity to the HTA process. Increased analytical sophistication could result in making assessment findings less understandable and transparent and therefore more exposed to resistance by decisions makers and other stakeholders. A recent commentary (Raftery 2010b) on the English NHS multiple sclerosis risk-sharing scheme highlights potential issues of governance, methodological requirements and ethics raised by these arrangements.

Implementation

The implementation of HTA-based decisions has assumed greater importance over the last decade. Although implementation was originally beyond the initial remit of almost all four bodies (Chalkidou et al. 2009), most have subsequently strengthened their focus on this issue, with NICE taking the most comprehensive approach⁷⁷. Since 2004, NICE has operated an implementation programme to support guidance adoption and to evaluate the uptake of guidance. Additionally, financial and regulatory levers are employed in England to promote adoption (e.g. funding mandate that creates an entitlement for patients to access technologies receiving a positive NICE recommendation) and strengthen links between guidance

⁷⁷ In Germany, the G-BA is responsible for implementation, not IQWiG.

and performance standards and payment systems; financial incentives are also used in Sweden.

Evaluations of impact

The rise of and investment in HTA in Europe, and in the four countries in particular, raises the question whether and how it has impacted policy and practice. As Jacob and McGregor (1997) note, ‘however excellent an HTA may be, if it fails to influence the workings of the health care system, it is without impact and must be considered without value’. In this discussion, we examine the direct impact of HTA on the following dimensions: (1) whether assessments are reflected in decisions or linked with policy; (2) if policy is adopted and integrated into clinical practice; and (3) the extent to which changes in practice result in better outcomes, in terms of health and/or budget impact. We also discuss its indirect impact on innovation.

Influence of coverage and reimbursement decisions

The impact of HTA on national policy varies across and within countries, but is arguably a function of the regulatory and legislative authority and instruments available to each HTA body. As intimated earlier, HTA has been most directly influential on national policy in England and Sweden, particularly with regards to pharmaceutical coverage and reimbursement. For example, as previously mentioned, in the case of NICE, the NHS Constitution makes all positive NICE decisions an entitlement, and in Sweden, a positive approval decision by the TLV is required for a pharmaceutical to be made available on the health system. In Germany, the GBA is not mandated to follow IQWiG’s recommendations, although it is required to provide justification if its policies deviate from the Institute’s advisement. This is also true in the French context, where HAS only gives an opinion to the Ministry of Health. However, available evidence suggest more than 95% of HAS recommendations are followed (Rochaix and Xerri 2009). Moreover, in the German case, under law, new drugs and inpatient medical services are covered by default and are assessed for possible exclusion only if the GBA requests an evaluation by IQWiG. Consequently, Germany pays higher prices and covers more new drugs than other European countries (Nasser and Sawicki 2009). However, the 2011 reform

strived to address this issue by requiring early comparative benefit assessment to prevent additional cost and harm of medical services without proof of benefit.

For countries with some level of decentralised decision making regarding access to or funding of treatments, there is concern that HTA has had less of an impact, or an inconsistent impact, on policy making at the local level. For example, as NICE only examines a fraction of all services provided across the NHS, most local decisions are not supported by HTA. Moreover, given funding mandates for local implementation of NICE technology appraisal guidance, concerns have been raised that implementation requirements potentially shift resources away from other, more cost-effective services (Devlin et al. 2003; Sheldon et al. 2004), although it is not clear what these services may be, as the opportunity costs of adopting particular technologies are not considered in assessments and local authorities rarely have formal mechanisms in place or the resources to make such judgements (Appleby et al. 2009; Audit Commission, 2005; Chisholm et al. 2009). In the Swedish context, Anell and Persson (2005) indicate that it is not clear the extent to which formulary committees, organised by the county councils, consider available HTA-generated recommendations, particularly economic evidence, to support decision making, and that the recommendations of the committees and the TLV often differ. Furthermore, HTAs may be limited in accounting for important geographic specificities in local policy making, in that use of a technology may be cost-effective in one region, but not in another due to different clinical practice patterns or patient population characteristics and needs.

The impact of HTA on policy is also dependent on the particular circumstances surrounding an assessment or decision. For instance, decision makers may not fully understand or accept an assessment, resulting evidence, or final recommendation, and even if they do, the consequences (e.g. a negative recommendation to adopt a given technology) may not always be accepted, especially if it results in challenge from industry, clinicians, patients and the general public. In Germany, for example, the Federal Constitutional Court challenged an earlier decision that restricted reimbursement for bioresonance therapy following request for treatment from a 19-year-old patient suffering from Duchenne muscular dystrophy. Based on the life-threatening nature of the illness, sickness funds were required to cover the treatment

despite a lack of solid scientific basis demonstrating benefit (Schmidt and Kreis 2009). Similarly in England, various stakeholder groups have pressured policy makers to circumvent negative NICE decisions on cancer drugs and multiple sclerosis treatments, and as a potential by-product, the new coalition government announced intentions to create a ‘cancer fund’ to pay for oncology therapies rejected by NICE.

Williams and Bryan (2007) suggest the more high profile the decision body or the decision being taken, the more likely external stakeholder groups and the media are likely to intervene if decisions are viewed as negative. For example, HTA decisions can significantly impact potential financial gains (or losses) for manufacturers, so they understandably exercise voice to ensure such interests are protected, especially if the decision of a particular HTA body has some degree of international influence, as in the case of NICE. Besides political considerations, assessments may or may not be influential in decisions depending on administrative capacities, equity concerns, broader societal preferences, and decision makers own values and experiences (Goddard et al. 2006; Owen-Smith et al. 2010). Furthermore, the evidence generated by HTAs may not sufficiently meet the information needs of decision makers or may not be made available early enough to benefit decisions (Hutton et al. 2006; Williams and Bryan 2007).

Clinical practice

The available evidence evaluating the influence of HTA on clinical practice is sparse and predominately relates to NICE. Several studies concluded that practice generally reflected the recommendations of the NICE technology appraisal(s) evaluated (National Cancer Director 2006; NICE 2006a, 2006b), but generally more so in the case of pharmaceuticals than procedures and devices (Sheldon et al. 2004). Other researchers, however, found limited impact on practice patterns and evidence of a high degree of regional variation in guidance adoption (Brickwood 2004; Hitchen 2008; Owen-Smith et al. 2010).⁷⁸ Yet, for some interventions, particularly cancer and obesity drugs, there is evidence of improved uptake of NICE-approved

⁷⁸ It is important to note that some of these studies were conducted prior to making implementation of technology appraisals mandatory and recent initiatives to improve guidance uptake.

therapies and reduced regional variation in care (Department of Health 2009; Sheldon et al, 2004). In a qualitative analysis of stakeholder perspectives on the usefulness of NICE guidance in practice (Owen-Smith et al. 2010), clinicians generally agreed with the importance of evidence-based medicine and found guidance useful, but its utility declined in certain contexts: the secondary care setting, when requisite funding was insufficient, and when they disagreed with the recommendations. Interestingly, several practitioners noted that the existence of NICE guidance can make it more difficult to resist patient demands, resulting in overtreatment, but that when the Institute said ‘no’ or placed restrictions on access, it offered a ‘good defence’ against blame for rationing treatment.

Similar conclusions regarding impacts on practice can be drawn in the French context. Criticisms have been raised that the process of disseminating guidelines is not well-structured and that adherence amongst physicians is poor, even though they are legally required to follow them in most cases. Moreover, guidelines are frequently produced on clinical issues that are not well-defined (i.e. little clinical certainty exists) and do not reflect areas where the most significant practice variations exist. Commentators have argued for better tools to prioritise and update or revise guidelines in attempts to be more responsive to emerging topics and stakeholder needs (Caniard 2002). The influence of HTA on clinical practice depends partly on available incentives for guidance adoption, in addition to strong professional support and a clear and robust evidence base demonstrating a given treatment provides value (Sheldon et al. 2004). Several countries have therefore introduced a range of mechanisms to support practitioner uptake, including the use of financial incentives, such as provider performance standards, and through educational and training programmes to enhance understanding of HTAs and how to access and interpret associated evidence. For example, the newly proposed Best Practice Tariff in England aims to align provider payment with NICE clinical guidelines by providing additional payment for delivering services that meet quality standards. To enhance professional support, a network of local experts was developed in Sweden help to ensure HTAs are applied in clinical practice (Sorenson et al. 2008a).

Health and economic outcomes

There is limited evidence demonstrating the health and/or budget impact attributable to changes in policy and practices associated with HTA. Some proponents of HTA promulgate its use based on conclusions that it reduces health-care costs. However, cost containment or cost-savings was never an objective of any of the four national bodies (Chalkidou et al. 2009), with aims of better quality of care, equitable access to care, and value for money taking precedence. In the case of NICE, its guidance has most likely been cost-increasing, in the order of £1.65 billion per year in additional NHS investment (NICE 2009). This is not surprising since most interventions that are deemed cost-effective are more expensive than their comparator interventions. The French HAS also claims that any adoption of cost-effectiveness analysis would not be used to save money (by reducing services), but attain more efficient use of resources (Chalkidou et al. 2009).

Whether overall costs increase or decrease as a result of HTA and how this should be evaluated depends on a number of factors. One issue relates to the baseline or starting point from which to derive conclusions about cost impact. For example, in France, where the use of drugs is significantly higher per capita than in England or Sweden, HTA may result in cost-savings without harming health outcomes. Another consideration is a particular system's structure, including its funding model. In England where budgets are set by government rather than supply and demand, 2002–2010 saw a significant increase in NHS spending, which NICE took advantage of, rather than caused, by then allocating some of the additional funding toward new technologies; a situation that will change in the future with restricted spending. In addition, the effect of HTA decisions on total costs is rarely measured or taken into account; therefore it is difficult to accurately assess impact on this dimension. NICE has tried to address this issue recently by costing the national impact of implementing its guidance for certain technologies or conditions. Finally, it is not entirely possible to assess the cost impact of the counterfactual; that is, if a given HTA decision had not been implemented.

Critics of HTA often cite potential negative impacts on innovation by creating yet more ‘hurdles’ for industry. However, there is no evidence coming from any of these countries (or elsewhere) to substantiate such concerns. In fact, as Chalkidou et al. (2009) point out, HTA systems can serve to create a more predictable and consistent way for industry to get their products to patients and for payers to make purchasing decisions. For products of true value, the HTA process should be of no real concern and even welcomed. In particular, for countries applying value-based approaches, manufacturers of truly innovative technologies will be rewarded with high or higher prices and from the onset will be armed with some indication of the type of innovation that is valued by decision makers (Claxton 2009; Hughes 2008; Kennedy 2009).

Conclusions

While HTA in Europe has evolved differently across jurisdictions over the last 10 years, it is evident that countries have generally strived to modify their methods and practices to improve the impact of assessments on policy and practice, meet national objectives and the various needs of stakeholders, and achieve greater transparency, legitimacy and relevance. Based on the last 10 years, it seems likely that countries will continue to evolve and improve upon their HTA processes in a variety of ways.

First, in order to better link decision makers’ needs with research agendas and account for the growing sophistication of health technologies, there will be increased attention placed on developing and considering different kinds of evidence (i.e. apart from evidence from traditional clinical studies) and ways to generate such information, such as observational research and new methods of evidence synthesis. This will also include new methods to evaluate different types of treatments, such as medical devices. In tandem, it is likely countries will continue to employ strategies such as CED and patient access schemes in order to reduce the uncertainty in decision making, while supporting access to potentially valuable interventions. Given

there is limited evidence on the effectiveness or cost-effectiveness of these strategies, evaluation is warranted.

Second, while HTA is often applied to coverage and reimbursement decisions, its influence in other areas of policy making seems set to grow. For example, the concept of value-based pricing (VBP) has gained traction. VBP is a method of setting prices for health technologies based on measured benefits to patients, in attempts to yield greater efficiency and to create a stronger link between evidence-based reimbursement and pricing decisions. A few countries (Sweden, Australia) have adopted variants of VBP, with the UK slated to follow by 2013–2014 (Department of Health 2010; Sweden Dental and Pharmaceutical Benefits Agency 2010). Evidence from Sweden suggests that VBP has resulted in reduced expenditure on drugs (Persson 2012) and higher profit margins for drugs that offer significant advances in therapy (Lundkvist 2002; Roughead et al. 2007), which may encourage industry to focus their research and development efforts away from “me-too” drugs and towards those that provide added-value. However, Sweden’s experience with VBP has raised implementation challenges, namely that reimbursement prices linked to the product and not the indication for a drug has led to increased variation in prescribing (Moise and Docteur 2007). Other potential areas of HTA application include developing (and updating) performance/outcome measures and payment tariffs. For example, the Quality and Outcomes Framework in the English NHS uses clinical indicators derived from NICE guidance to assess provider performance. Similarly, financial incentives are used to reward or penalise providers that do not adhere to NICE guidance. There is risk, however, that expanding the uses of HTA ignores its limitations and expects too much of the approach.

Third, discussion around the disinvestment of existing, ineffective interventions is mounting to achieve greater health system efficiency. Some countries (e.g. England, France, Sweden) presently require re-assessments of technologies (or entire therapeutic categories) after they have been on the market and used in routine practice for a specified number of years, and use such evidence to modify reimbursement and pricing status, or to remove interventions from the benefit package altogether. Given the current economic situation and limited budgets, the need for disinvestment of low-value interventions will grow in importance. The

practical implementation of disinvestment decisions, however, is marked with controversy and problems, which must first be addressed. Transparent, formal, and evidence-based approaches to identify ineffective and low clinical value treatments are needed (Flynn and Gericke 2012).

Fourth, in light of existing hurdles to use or implement HTA at the local level, a potential area for future work is how to better localise or contextualise evidence-based decision making. Potential strategies could come in the form of more effective financial planning tools or implementation of horizon scanning or alternative systems to inform decision makers of new technologies that may be introduced. Other mechanisms might include guidance and training to support local decisions on interventions not assessed at the national level. For example, the use of a ‘mini-HTA’ tool is currently in use in hospitals in Denmark to help guide purchasing decisions. With GPs now having greater responsibility over NHS’s budget in England, ‘mini-NICE’ programmes may also be set up to support Clinical Commissioning Groups.

Fifth, international collaboration between HTA bodies is set to grow. While general consensus exists that the appraisal process should be undertaken within national and local contexts, there are potential efficiencies to be gained from enhanced collaboration around assessments. Increased sharing of information (e.g. methods, data requirements, results) across countries may save costs and reduce duplication. International collaboration may also facilitate evidence development for promising technologies, where existing data are often limited and pooled expertise is increasingly required. The feasibility and effectiveness of international collaboration is dependent, however, on addressing potential challenges, such as attaining agreement on review priorities and assessment perspectives (e.g. societal vs. payer), standardising methods, ensuring that supporting studies or assessment meet the needs and circumstances of different countries, and protecting the confidentiality of commercial data. The EUnetHTA is working to address some of these hurdles.

Finally, the lack of studies on the impact of HTA constitutes an important gap in understanding the role and influence of HTA in health policy. As one commentator noted, whereas the previous 10 years have been well-spent on building

the HTA infrastructure and evidence base, the next decade should focus on ascertaining outcomes (Straus 2004).

Study 6: Decision making under uncertainty: An investigation of international coverage with evidence development policies in the context of medical devices

Introduction

Making coverage decisions on new health technologies is an enduring challenge for payers worldwide. Patient access to new technologies and support for innovation must be balanced against broader demands for prudent use of resources and increasing requirements by policy makers and payers that coverage determinations be based on solid evidence of clinical- and, in some cases, cost-effectiveness and wider socio-economic benefits. Often, however, the available evidence for individual technologies, especially with regards to routine (“real world”) use or compared to existing interventions, is suboptimal or inconclusive at the time coverage is determined (Claxton et al. 2012; Sorenson et al. 2012).

Determining coverage policy in cases where evidence is uncertain or limited can have important consequences. For instance, coverage may be denied for potentially beneficial technologies or delayed, which may hinder patient access to important new treatments. Conversely, technologies may be covered and integrated into clinical practice that later prove ineffective or lacking in value for money. This could displace resources that could be availed to pay for more effective treatments, potentially resulting in suboptimal patient care. Additional costs may also be incurred from having to reverse an inappropriate coverage decision (Eckermann and Willan 2008; Palmer and Smith 2000). Moreover, without sufficient evidence about the risks and benefits of existing treatment options, payers, physicians, and patients may lack the information necessary to guide the best care decisions.

To address these issues, several jurisdictions in North America and Europe have established CED policies, which provide provisional coverage for a promising, but unproven, intervention, on the condition that additional data are generated to inform coverage and payment policy (Tunis and Pearson 2006; Hutton et al. 2007; Mohr and Tunis 2010). Upon completion of a CED study, if the findings substantiate

that the new treatment is better than existing options, then payers may expand coverage to more beneficiaries or decide to cover the treatment permanently. Conversely, payers may deny coverage or restrict use to certain patient subgroups.

CED schemes assume different names across jurisdictions, such as ‘only in research’, ‘field evaluations’, ‘conditional coverage or reimbursement’, or ‘access with evidence development’. However, a common feature is that they embody a systematic approach to collecting data on outcomes in regular clinical practice – evidence that is subsequently used to inform or modify coverage and/or reimbursement decisions.

Despite the growing interest in and use of CED internationally, much of the published evidence on the approach has been conceptual in nature. A few case studies on a single country or particular technology have been published (Briggs et al. 2010; Carino 2006; Chalkdou et al. 2007; Claxton et al. 2010; Dhalla et al. 2009; Levin et al. 2011; Longworth et al. 2013; Ramsey and Sullivan 2005; Tunis and Pearson 2006; Whicher et al. 2009), but none that compare the use of CED in different countries and in the context of medical devices. Devices, in particular, may be viable candidates for the CED approach, given a range of challenges associated with generating robust evidence on their benefits and costs (Ciani et al. 2013; Drummond et al. 2009; Sorenson et al. 2011). For instance, there are particular characteristics of devices that make it difficult to conduct RCTs, especially for first generation technologies. Unlike pharmaceuticals, it is difficult to control (or ‘blind’) the treatment assignment, which introduces the possibility of bias in the outcome assessment. Moreover, there appears to be some concern that CED, while attractive in principle, has not resulted in expected outcomes or uptake by payers, which suggests a need to better understand associated challenges with applying the policy in practice.

This paper aims to meet these existing evidence gaps and is structured as follows. First, the study methods outlined. Second, a brief overview of international CED approaches is provided, including select case studies of CED applied to devices. Third, key challenges for CED policies in general and for devices in

particular are identified and discussed. Finally, the paper concludes by outlining some key areas for improving the use of CED in practice.

Methods

A two-phased approach was used to address the research aims. First, a literature review on international CED schemes was conducted. For purposes of guiding the study, we defined CED as: *a form of conditional reimbursement characterized by restricted coverage for patients enrolled in a study that is designed to collect better data around the safety and effectiveness of a medical technology*. The review analysis focused on CED programs in Europe and North America, namely Canada, France, Germany, the Netherlands, Switzerland, the UK and US. While other countries, such as Sweden, also operate national CED programs, they have not been applied to devices to date, only pharmaceuticals. Therefore, they were excluded from the study. The countries selected represent a mix of health care systems (public, private, mix) and countries with more established and new CED schemes. In addition to specific national CED schemes, the available literature on cases where CED has been applied to medical devices was searched as well as on the CED approach more generally. Second, semi-structured, in-depth expert interviews were conducted to supplement the information gathered from the literature and to better understand some of the key opportunities and challenges faced to date by the respective national experiences with CED, particularly with regards to medical devices.

For the literature review, several key databases were searched, including PubMed, EMBASE, MEDLINE, Scopus, and Google Scholar, using the following search terms: “medical technology”, “medical devices”, “coverage with evidence development”, “access with evidence development”, “conditional reimbursement”, “innovative reimbursement”, and “restricted reimbursement”. The search also included grey data sources, such as relevant website materials, policy documents, and academic working papers. After the abstracts were identified (total of 75) and reviewed for relevance, the full published papers or materials were obtained and reviewed. In total, 50 articles were gathered and reviewed. Appendix A contains the list of articles reviewed. Besides providing information on international CED

policies, the available literature informed the development of semi-structured interview guides used to conduct the expert interviews. Appendix B contains the interview guides used in the study.

The study sample encompassed experts from three different groups – payers/HTA bodies, the medical device industry, and academia/policy analysts – which allowed for a diversity of opinions of those involved in CED to be captured. The sample of potential interview participants was developed based on the authors' knowledge of and experience with experts in CED and related areas, and through snowballing techniques to identify other possible informants. In particular, the sample represented individuals involved in the development and operation of CED schemes, those heading reimbursement and pricing departments within industry, and academics with specialisation in CED, HTA, evidence-based policy making, and related fields of study.

Potential participants were sent an email inviting them to participate in an interview and were provided a summary description of the study. A total of 27 experts were invited to participate, of which 22 agreed to be interviewed (81% response rate). Appendix C provides an overview of the interview sample. All interviews were conducted by telephone by both members of the research team and were approximately 45 to 60 minutes in length. An interview response coding guide was used to aid collection and later analysis of the interview data. Expert participants were not remunerated for their involvement, but were offered a copy of the final study findings upon completion of the research.

Upon completion of the interviews, the interview data was synthesised in two main ways. First, the data were used to supplement the information gathered in the literature review, particularly to understand the different national approaches to CED and the cases where devices have undergone CED. Second, the data were used to gather expert opinion on CED. The interviews were read and notes were taken and key themes and word usages were extracted and imputed. Illustrative quotes were gathered to highlight and support key points associated with the identified themes. Interviewees were promised anonymity and, as a result, they are only identified by

their role in the process and the country where they work. The following section presents the study findings.

Results

Overview of international approaches to CED

Countries in Europe and elsewhere have adopted CED policies to better inform coverage policies for new device and other technologies (Table 12). Box 3 outlines select cases of medical devices CED schemes in different jurisdictions, and Appendix D provides a broader range of examples of CED schemes in Canada, UK, and the US.

While some policies have been in operation for a decade or longer (e.g. Canada, Switzerland, the UK), others are just getting started or are still under development, such as in Germany and the Netherlands. Countries generally share similar aims for CED, namely to address outstanding uncertainty about the benefits and/or costs of new technologies; allow more flexible coverage decisions (as opposed to either yes or no); and, enhance the existing evidence base on new technologies. Those jurisdictions with culture traditions for supporting innovation and industrial policies toward this end, such as France and Germany, also focus on the benefit of CED to speed patient access to new therapies and to support industry. However, countries have adopted divergent approaches to CED, from involving different stakeholders groups to requiring different criteria for which technologies and interventions are eligible for CED and who oversees and runs studies.

For example, while all countries' CED policies involve the regulatory or advisory body directly responsible for making coverage and reimbursement decisions, the jurisdictions differ with regards to other stakeholders involved in the CED process. Canada, the Netherlands, UK, and, occasionally, the US involve their national health or clinical health organisation in advising and/or overseeing CED studies. While still in the planning phase, Germany and the Netherlands intend to involve physician, hospital, and insurance associations in the process (which Switzerland already does) by allowing these groups to submit applications for CED

candidates. The role of manufacturers also differs across countries, with the most central role likely in Germany and perhaps in Switzerland. In Germany, for example, manufacturers can submit applications for CED and, if selected, are involved in negotiating the terms of the studies and funding them. In other countries, such as Canada, industry plays a minor to no role in CED schemes.

The funding of CED studies also varies between countries. In Canada, Germany, and the Netherlands, public bodies – either the national payer or health research organisation(s) – fund CED studies, although in Germany manufacturers are sometimes be responsible for financing studies. In these countries, funding processes are also more formalised, where funding is guaranteed if a particular CED scheme is selected to go forward. Funding for studies in the UK and US, for example, is not required or pre-determined *a priori* before a CED recommendation is given. Switzerland is similar, although it is generally the applicant who is responsible for funding the study if it is indeed conducted.

Finally, countries differ slightly in terms of preferences for, or experiences with, particular types of studies. The newer CED programs, such as in Germany and the Netherlands, have expressed interest in a variety of study methods (e.g. RCTs, registries, prospective case series), while countries with more experience with CED (e.g. France, US) have traditionally leaned toward the use of RCTs (with the exception of Switzerland), although this could be due, in part, to the types of technologies involved in their respective CED schemes to date.

Table 12: *Comparative overview of international CED programs*

Country	Name of Program	Aims	Year Established	Technologies Included	Actors Involved	Procedures and Methods	Funding Sources	Examples of CED
Canada	Conditionally-funded Field Evaluations	Assess real world performance; address outstanding uncertainty about benefits/costs; improve coverage decision making.	2003	Non-drug technologies (devices and procedures)	OHTAC, PATH, THETA, ICES, Ontario Health Ministry	<ul style="list-style-type: none"> Decision makers request field evaluations. OHTAC allocates evaluation to affiliated academic institutions. All types of study designs accepted. To date, majority of evaluations have entailed observational or registry studies. 	OHTAC funds the evaluations; Ministry funds device (or procedure) if not yet insured.	<p>Over 40 studies to date.</p> <p>Examples include:</p> <ul style="list-style-type: none"> PET DES CT angiography Sleep apnea device
France	CED	Avoid delays in coverage and utilization of innovative technologies.	2011	Devices, procedures, and procedures involving devices.	Ministry of Health, HAS	<ul style="list-style-type: none"> To date, no clear eligibility criteria or methodological standards (case-by-case basis). HAS generally prefers RCTs, although registries have been established. 	Information not available.	<p>Several studies are underway, including:</p> <ul style="list-style-type: none"> TAVI Retinal implants CT angiography Monitors for continuous blood glucose measurement
Germany	CED	Address lack of sufficient evidence on new technologies; allow more flexible coverage decisions; link clinical researchers and decision makers; enhance transparency; and, strengthen history of support for industry and innovation.	2012	Procedures and procedures involving devices.	Ministry of Health, GBA, sickness funds, German hospital association, Association of Board Certified Physicians, IQWiG.	<ul style="list-style-type: none"> GBA, physician and hospital associations, and (now) manufacturers can submit applications to assess a procedure for reimbursement. If procedure deemed to offer “potential”, GBA, in collaboration with IQWiG, determines if eligible for CED. If so, a directive requesting a study is published and terms of the study negotiated between the GBA, manufacturer, and other parties involved in conducting the study. GBA has expressed interest in 	If application submitted by manufacturer, they must fund study; otherwise, GBA funds.	None to date. First studies anticipated in 2014/2015.

						a wide range of studies (RCTs, non-randomized studies, comparative studies).		
Netherlands	Conditional Entry	Collect evidence on new (not yet part of the benefit package) interventions, while providing patient access.	2012	Procedures, devices, and drugs.	CVZ, NZa, ZonMw	<ul style="list-style-type: none"> • Application for conditional entry made by a patient or provider association. • Candidates for conditional entry selected based on well-defined evidence gaps; quality of research protocol; feasibility of collecting relevant evidence; value of evidence. • Currently, no specific guidance on what study designs are acceptable, but must be “methodologically acceptable and realistic”. • To date, majority of studies have been large, multi-center RCTs. 	ZonMw funds all conditional entry studies.	<p>A number of CED studies are now underway, including:</p> <ul style="list-style-type: none"> • Radio frequency denervation for chronic low back pain • Renal denervation for therapy resistant hypertension • Intra-arterial thrombolysis for acute stroke
Switzerland	CED (yes, in evaluation)	Provide temporary coverage for novel and promising interventions, while additional evidence is gathered.	1996	Procedures, procedures involving devices, and drugs.	SFOPH	<ul style="list-style-type: none"> • New technologies: SFOPH performs horizontal scanning to identify new interventions that lack sufficient evidence of effectiveness, efficiency, and appropriateness. If manufacturer seeks coverage, a CED arrangement may be requested/imposed. • Old technologies: If questions arise regarding effectiveness, a request can be made to SFOPH to reclassify from covered without evaluation to CED. • In either case, CED determined on a case-by-case basis. • Use of prospective multi-centre case series observational studies most common. 	Costs of studies covered by manufacturers only.	<p>Around 27 studies to date: 10 diagnostic procedures (7 involving devices); 10 alternative medicines; and, 7 surgical interventions (5 involving devices).</p> <p>Examples include:</p> <ul style="list-style-type: none"> • PET • Total disk replacement • Balloon kyphoplasty • TAVI

UK	Only in Research (with limited use of Approval with Research)	Provide coverage to promising interventions not yet supported by sufficiently robust evidence, while additional data is collected.	1999	Procedures, devices, and drugs.	NICE, NIHR	<ul style="list-style-type: none"> • Eligible technologies determined on a case-by-case basis, as a result of the technology appraisal process. • NIHR systematically reviews feasibility and priority of potential study. • NICE outlines broad areas of inquiry the research should address. • Either NIHR or the manufacturer initiates study. • Use of RCTs and registries most common. 	No standard or requirements for funding. NIHR or manufacturer may fund study.	<p>Over 25 studies to date.</p> <p>Examples include:</p> <ul style="list-style-type: none"> • PET • ICDs • Metal-on-metal hip implants • Drainage, irrigation and fibrinolytic therapy (DIFT) • Laparoscopic surgery
US	CED	Allow greater flexibility in coverage determinations; link coverage to efforts to generate evidence needed to gain greater certainty on the benefits and harms of particular technologies.	2006*	Procedures, devices, and drugs.	CMS	<ul style="list-style-type: none"> • Through the NCD process, CMS determines candidate technologies for CED process. • CMS published requirements for study. • Interested parties submit study proposals for approval. • Studies generally overseen by party who submitted proposal. • Use of RCTs and observational studies most common. 	No standard or requirements for funding. Public agencies, such as NIH or AHRQ may fund studies, as well as manufacturers, medical associations, or academic research groups.	<p>Over 15 studies to date.</p> <p>Examples include:</p> <ul style="list-style-type: none"> • PET • ICDs • Lung Volume Reduction Surgery • Angioplasty and stenting • Transcutaneous Electrical Nerve Stimulation

*Applied since 1995, but no clear policy until 2006.

Source: Authors' compilation based on literature review and expert interviews.

Box 3: *Select case studies of medical devices CED schemes*

Drug-Eluting Stents (DES) - Canada

In 2002, the Ontario Medical Advisory Secretariat (MAS) completed a secondary literature-based technology assessment of the clinical effectiveness of DES compared with bare-metal stents (BMS). It concluded that RCT evidence would likely demonstrate that DES was more effective than BMS, after which there would be a steep uptake for DES. However, when the initial RCT results on DES were published later that year, there was uncertainty regarding the generalizability of the results. Accordingly, OHTAC recommended that the Health Ministry commission a field evaluation from PATH. PATH proposed a prospective observational study, which took advantage of both an existing province-wide registry established by the Cardiac Care Network (CCN) of Ontario and the ability to link this registry to administrative databases housed at the ICES. Faced with rapid growth in the use of stents and projections for a major shift away from bare metal to the more expensive DES, the Ministry agreed to cover DES only if additional data were collected on the effect of stent choice on outcomes. Consequently, new fields were added to a pre-existing CCN database to facilitate a study comparing different stent designs. The aim of the study was to estimate the reduction in risk of revascularisation within two years of treatment with DES, compared with BMS, as well as comparative cost-effectiveness. Hospitals were able to provide DES free of charge to patients enrolled in the study (Bowen et al. 2007; Tu et al. 2007).

The study found that DES were effective in reducing target-vessel revascularisation among patients at highest risk for restenosis, but had no effects on death or myocardial infarction (Bowen et al. 2007; Tu et al. 2007). Based on this study, OHTAC recommended that DES be restricted to use among high-risk patients (those with diabetes, or particularly long or narrow lesions) to improve the appropriate use of DES. Estimates suggest that this controlled diffusion of DES led to estimated savings of \$35 to \$58 million (Bowen et al. 2007).

Stenting and Aggressive Medical Management – United States

In 2006, CMS issued a CED policy for intracranial stenting for preventing recurrent stroke in intracranial stenosis (CMS 2006). The policy provided coverage for high-risk Medicare beneficiaries only when they were enrolled in an FDA-approved trial for Category B Investigational Device Exemption study. Over the years, there has been rapid adoption of percutaneous transluminal angioplasty and stenting (PTAs) for prevention of a second stroke in high-risk patients. The trial evaluated if treatment with PTAs and medical management was superior to medical management alone in the treatment and prevention of a second stroke in high-risk patients. The study was conducted at approximately 50 sites in the US with financial support from the National Institute of Neurological Disorders and Stroke (NINDS) and other public and private sponsors. The CED scheme played an essential role in expediting enrolment in the Stenting and Aggressive Medical Management for Preventing Recurrent Stroke in Intracranial (SAMMPRIS) trial. The results of this study demonstrated that patients undergoing PTAS have a much higher rate of stroke or death (14.7%), as compared to patients receiving medical management alone (5.8%) (Chimowitz et al. 2011). Consequently, the trial was stopped early due to the high risk of early stroke in patients undergoing PTAs.

Total Disc Anthroplasty (TDA) – Switzerland

Short-term clinical results for TDA in the treatment of degenerative spinal diseases are promising, but its use has a relatively young history in clinical practice and long-term follow-up data are limited. Consequently, the Swiss SFOPH commissioned a nationwide prospective, multi-centre, observational study (SWISSpine) before making a final decision about the reimbursement of TDA (primarily for cervical and lumbar TDA and for balloon kyphoplasty). The national registry was established in collaboration with the implant manufacturers, the Swiss Spine Society, and the Institute for Evaluative Research in Medicine at the University of Bern. Funding of the registry was shared amongst the participating manufacturers.

Between March 2005 and June 2008, 1682 interventions (808 cervical, 427 lumbar, 331 balloon kyphoplasty) with implantation of discs from several different manufacturers were performed. Surgery, implant, and follow-up case report forms were administered by spinal surgeons. Co-morbidity questionnaires, EQ-5D, others forms were completed by patients. Data collection transpired pre- and peri-operatively at 3 months, 1 year follow-up, and annual thereafter. The three year study results suggested that both cervical and lumbar TDA are relatively safe and efficient procedures concerning pain reduction and improvement of quality of life (Diel et al. 2009; Schluessmann et al. 2009; Schluessmann et al. 2010). However, mid- to long-term effectiveness and safety were not established (and the patient sample was relatively small) and therefore the SFOPH required a 10 year follow-up and extension of the study before making a final coverage determination. Conversely, the SFOPH agreed to permanent coverage of balloon kyphoplasty following the study based on findings suggesting a significant and clinically relevant reduction of back pain, improvement of quality of life and preoperative segmental kyphosis, and reduction of pain killer consumption.

Laparoscopic Surgery – United Kingdom

NICE first reviewed the use of laparoscopic surgery for colorectal cancer in 2000. Based on the assessment, the Institute determined that there was a lack of evidence about the long-term outcomes of the laparoscopic approach compared to the conventional open technique (NICE 2006). As a result, NICE recommended that laparoscopic surgery for colorectal cancer only be used as part of a randomised controlled trial (Chalkidou 2007). This decision encouraged recruitment to the then ongoing UK-based Medical Research Council CLASICC trial, which has since provided the necessary evidence to support a revised recommendation. In particular, the trial indicated that the long-term outcomes for patients are equivalent for both techniques. In addition, there were important additional benefits associated with laparoscopic surgery, both in terms of shorter hospital stays and the ability of patients to return to normal activities post-operatively faster than with conventional surgery (NICE 2006). Therefore, NICE revised its recommendation to support use of the technique in routine HHS practice.

Key challenges for CED policies

While the various jurisdictions have developed their CED policies differently to meet the particular needs of their respective health care systems, the analysis suggests that they share many common challenges in ensuring the effectiveness and efficiency of CED. The following section discusses five key challenges facing CED policies identified in the interviews, including establishing a clear framework for initiating and overseeing CED studies; identifying and applying appropriate study methods; funding CED studies; incentivising research; and, applying new evidence to coverage policy.

Establishing a clear framework for initiating, overseeing, and stopping CED

CED initiatives are likely to be more credible to patients, policy makers and industry if there is a well-defined, transparent and consistent approach for initiating topics for CED. However, “current processes do not provide public stakeholders the level of clarity of specificity needed to ensure a transparent, predictable process” (*US policy analyst 1*) and operate on a “reactive case by case basis” (*UK policy maker 1*) with “no rules or standards” (*Switzerland policy analyst 1*), rather than having a well-defined priority-setting procedure and criteria to select technologies for this approach. This particular issue was highlighted in recent public comments on ways to improve the CED program in the US (CMS 2012a) and has been the focus of research in the UK on the circumstances within which NICE should recommend use of health technologies only in the context of further evidence development (Claxton et al. 2012). Identifying clear criteria for initiating CED studies for particular technologies was considered particularly important given the growing number of devices on the market. “There are lots of devices; we cannot have CED schemes for all of them” (*US policy analyst 2*)... “CED is good policy, but we do not want to make it the default option in all cases where uncertainty in the evidence exists” (*UK policy analyst 5*).

Informants provided specific ideas for when CED should be employed. “Some devices are low-risk with low-budget impact, which are probably not good candidates for CED, especially given the financial demands of CED studies....rather,

CED should be applied to those technologies that offer the potential for significant improvement in care, but there is some material uncertainty about effectiveness or other important endpoints” (*UK policy analyst 2*). To that end, the potential impact on patients was seen as important criteria for CED. ”Higher risk medical devices, such as implantables, are appropriate candidates for CED” (*Industry representative 3*) as are “devices that address a true unmet medical need” (*France policy maker 1*). A few informants suggested that CED schemes cover both new and existing technologies, although they also duly noted the challenges in doing so, including limited resources to conduct CED studies as well as “difficulties recruiting patients and collecting evidence once a product is already routinely used” (*Netherlands policy maker 1*). Respondents emphasised that CED is not the sole tool to address issues of uncertainty. In cases where there is uncertainty about the financial impact of a technology or whether it will be used appropriately in practice, risk-sharing agreements and performance-based contracts were deemed more suitable.

Issues around clarity and transparency were also evident with regards to planning and overseeing CED studies. Informants generally thought that insufficient attention is currently given, prior to initiating studies, to agreeing the key evidence gaps, the acceptable quality of evidence, study design and key endpoints, such as what constitutes meaningful clinical benefit, when to stop CED studies, and how studies will be funded. “In planning the scheme, it is important to be precise and clear about the evidence that is lacking and that such evidence is critical to the decision...CED cannot be used to simply delay a difficult decision” (*UK policy analyst 2*). The feasibility of additional evidence collection is central to such discussions. “For a scheme to work, it has to be feasible to collect the data through the necessary study design in the timescale required....time, resource, and financial burdens of additional data collection all need to be carefully considered” (*US policy analyst 2*). Informants carried these sentiments over to more procedural considerations, such as outlining clear roles for the parties involved, timelines for key tasks, and how collected evidence will be linked to the coverage decision.

Greater predictability in these areas was also considered particularly important by and for industry. “If we are going to use these policies, they need to be predictable...there needs to be clear rules of engagement in order to incentivise

manufacturers to play ball” (*Industry representative 3*). Therefore, the challenge is to have a “clear agreement in advance among the main parties involved on when to review the data and how to use the evidence, especially given that industry tends to want pre-set rules to ensure any additional evidence collection will directly support a reimbursement decision, while the authorities want less specification, in order to retain their decision flexibility” (*UK policy analyst 2*).

A lack of clear terms of stakeholder roles in CED also introduces challenges in terms of aligning policy makers, manufacturers, and the research community, especially as these different groups hold different expectations and reservations around CED in practice. For example, “industry may feel that CED gives payers license not to make a decision based on existing data, while payers contend it provides an excuse for manufacturers not to collect adequate data in the first place” (*US policy analyst 2*) and the “research community feels constrained by timelines, lack of adequate funding, and the often political nature of the CED process” (*UK policy maker 1*). Several of the industry representatives interviewed highlighted the importance of bringing industry to the table when planning CED studies. “Industry needs to be involved throughout the process, especially in determining appropriate evidence and designing studies. A trial of coronary stents in France did not involve industry, but rather was overseen by centers that had no experience with implantation or with stents in general. The design of the trial did not take into consideration issues of the learning curve with implanting stents. In the end, it was a mess and negatively impacted the market for coronary stents. CED cannot be done in isolation of the authorities” (*Industry representative 5*). Another example relates to a case of negative pressure wound therapy, where the “authorities wanted to see wound closure as the primary outcome, but to those developing the intervention, this was not an appropriate endpoint. Rather, the aim of the therapy is to first get skin rosy and then switch to cheaper, advanced wound therapies, keeping the wound moist to heal. Authorities required manufacturers to take this approach, which was a recipe for failure” (*Industry representative 3*). However, others, particularly researchers, maintained the industry-sponsored CED studies must be “vetted carefully” (*US policy analyst 2*).

Identifying and applying appropriate study methods

One of the key challenges of CED is designing and executing an appropriate study to address any outstanding gaps or uncertainties in the existing evidence. In some regards, devices introduce unique challenges to this process, as acknowledged by almost all informants. “Devices can be difficult to randomise against a no intervention arm and it is important to take into account the operator-dependent nature of certain devices and procedures and the associated learning curve (*Industry representative 3*). Operators can entail not only clinicians, but also patients. In addition, “some devices bring long-term benefits that require assessment of outcomes over the course of several years” (*France policy maker 1*). The need for long-term data is “especially important in the case of implants” (*Switzerland policy analyst 1*). When evaluating devices, it is also important to account for the fact that “their development is iterative in nature and that their effectiveness is sometimes contingent in part on organisational factors” (*Industry representative 2*). To that end, “studies for devices often require involvement of a greater number of professionals and organisations, as use tends to entail a greater number of actors and depends more heavily on the clinician or surgeon, compared to other technologies, such as drugs” (*Netherlands policy maker 1*). In addition, it can be “difficult to accurately capture device use in studies” (*Switzerland policy analyst 2*), as they generally do not possess a unique identification code that allows for identification and tracking. However, policies to implement a UDI system to track and monitor device use have been introduced recently in both Europe and the US, which may help improve data collection within CED studies.

Yet, there was some consensus among informants that the evaluative challenges faced by devices are not always sufficiently acknowledged in planning CED studies. “Policy makers do not fully understand the nuances and difficulties in studying devices. Similarly, industry, which is largely composed of small enterprises, is generally unaware of the difficulties and are desperate to obtain some level of coverage and access for their product...these pressures lead companies to underestimate the difficulty in conducting studies” (*UK policy analyst 2*).

Several informants noted challenges for CED in terms of accounting for the diversity of devices in designing appropriate studies, making it difficult to pre-determine certain standards or requirements for studies. “There are a lot of different devices with different technical specifications and they are often used in the same procedure, which makes it difficult to standardise” (*Germany policy maker 2*)...”different devices have different issues, which would impact study designs and data collection strategies” (*Industry representative 5*). Consequently, rather than determining a “gold standard”, selecting appropriate methods for device studies depends on the particular technology, as well as the evidence at the time of market authorisation, the outstanding evidence gaps, and available treatment alternatives, if any. “If the uncertainty regards treatment effect or comparative effectiveness, it is difficult to get away from doing an RCT”, while if the uncertainty is about long-term effectiveness or safety, then a well-coordinated registry or large, prospective case study would be indicated” (*UK policy maker 2*). The type of study may also depend on the stage of development of a device. For example, “if a second generation device is coming in the next couple of years, it will not make sense to conduct an RCT” (*Industry representative 3*).

Some differences in opinion emerged with regards to the challenges in conducting RCTs for devices, reflecting recent debates among analysts in the field (Drummond et al., 2009; Taylor and Iglesias, 2009). One perspective is that RCTs for devices are not always possible and therefore “evidence might come from registries and observation trials” (*Industry representative 1*). Another perspective, most often held by policy makers, is a preference for RCTs comparing a new device with the current best treatment standard...”[we] do not accept that devices cannot be effectively studied via RCTs” (*Germany policy maker 2*). A few informants suggested that it is not a question of whether RCTs are possible, but if they are designed appropriately. “RCTs are still possible, if one designs the trial appropriately and to the specificities of the technology. Trials should be designed for the technology, not technologies for certain types of studies” (*Industry representative 5*). Considerations regarding a robust study design also apply to other types of studies, such as registries. Ideally, registries should be “representative of hospitals (or other health care setting) nationally, include a comparator, and collect high quality evidence” (*Switzerland policy analyst 1*). Overall, however, it was generally believed

that it is most important to adopt the most feasible and practical methodological approach from both a cost and time perspective....”we need to be pragmatic” (*UK policy analyst 3*).

Funding CED studies

The costs associated with CED medical device studies are considered by informants to be one of the most important challenges to effective and impactful CED programs. Informants recognised a myriad of associated costs, including the costs of designing a study; recruiting patients and, in some cases, training physicians; aligning organisational needs (e.g. additional equipment requirements); and, overseeing studies and subsequent follow-up. Other indirect costs were also recognised, such as the transaction costs of reaching agreement on the terms of a CED study with the manufacturer, costs for new systems to track or monitor studies, and the cost of delays in approving a potentially beneficial technology for widespread use.

Such cost considerations have limited the use of CED in some countries, deterred buy-in from key constituents, and impacted the robustness and quality of the studies themselves. As highlighted by one informant, “the cost of CED studies can hinder buy-in from involved parties...they see the costs of funding CED studies, but not the benefits in practice” (*Canada policy maker 1*). “Lots of device companies are SMEs, especially given the current economic situation, and do not have the infrastructure and resources to conduct studies, which can negatively impact not only the ability to conduct studies, but the quality of the studies that are conducted (*UK policy analyst 3*). Funding of CED studies may be particularly challenging for procedures that involve devices, as there are typically several different manufacturers of the device involved.

The constraints on funding CED studies by public authorities (e.g. payers, HTA agencies, other research organisations) further hinder the success of CED schemes. As previously discussed, agencies such as the CMS in the US and NICE in the UK do not possess designated research budgets to fund CED studies and funding through other public agencies, such as NIH or NIHR, is tenable at best, especially

under tight economic conditions. “If the relevant policy maker does not hold the purse strings, it is difficult to mandate studies” (*Industry representative 3*). Coupled with the limited ability of or incentives for device manufacturers to fund studies, there is the inherent risk that CED will never get off the ground or be completed properly.

Incentivising research

Several informants mentioned issues around insufficient or misaligned incentives to conduct CED studies. In the Netherlands, one of the “biggest problems has been getting commitment from both physician and manufacturers to collect the necessary data...some of the evidence dossiers submitted at the end of studies have been less than impressive, which is one of the reasons that the length of CED studies was extended from three to four years in the recent revision of the scheme” (*Netherlands policy maker 1*). The lack of incentives for CED evidence generation relates to issues associated with the CED approach and process itself, inherent characteristics of a given health care system, and particular characteristics of the device industry.

To the former, “given that many CED studies are not mandatory for manufacturers or clinicians, it can and has led to limited participation” (*UK policy maker 1*) and “it is not always to the advantage of physicians to conduct studies and recruit patients” (*Netherlands policy maker 1*), especially if “they get paid, regardless of participation in the study” (*Switzerland policy analyst 2*). Moreover, although informants recognised the benefit of providing coverage while additional evidence is collected, they also noted that allowing patients early access to a technology may affect the prospect of the research being conducted. Manufacturers may have less incentive to invest in additional research about a technology once it is covered; physicians might consider further clinical trials or other studies to be unnecessary and unethical; and, patients might be unwilling to participate if they already have access to the new technology. Physicians may also not understand the importance or necessity of collecting the requisite data, which is one of the factors that have hampered the SWISSpine registry discussed previously.

Informants also believe the lack of a clear framework or guidance for CED disincentivises commitment from stakeholders to move forward on potential CED studies. “The process for CED is unclear...what is the incentive for industry to conduct additional research? If manufacturers undertake additional R&D they need some direction that the results will inform policy...and when” (*Industry representative 5*). “The role of evidence in policy making is not clear or predictable, which sends wrong signals to manufacturers to collect the necessary evidence” (*UK policy maker 3*). The time and effort it takes to agree a CED scheme can act as a deterrent by “delaying access and hindering interest in the approach” (*Germany policy maker 1*). In relation, in order for CED studies to be conducted successfully, “readily available and integrated data collection and monitoring systems” (*Switzerland policy analyst 1*) may be required. Some providers or hospitals (and other relevant parties involved in the study) may not be equipped with such systems, which may deter the timely completion of CED studies or the collection of high quality data.

A few informants mentioned the fact that some CED schemes, such as Germany, only consider evidence and studies conducted in that particular country context, which was viewed as a limitation to the CED approach that should be addressed. “If the German trial results are inconclusive, there is a risk that those involved might want to wait until trials in the other countries, if any, are complete, which will just extend timelines and there is the cost of waiting” and “manufacturers are increasingly global and therefore desire to conduct trials on an international basis; restricting studies to German borders could hamper the studies associated with CED and threaten the German CED approach (*Germany policy maker 1*). Restrictions on national CED studies also have implications on the generalisability of the generated evidence....”why invest in these studies when the results will only be applicable to a certain country context” (*Industry representative 4*).

Several informants noted the existence of system-wide disincentives that could potentially threaten the effectiveness of CED in their respective jurisdictions. For example, in the Netherlands, “temporary reimbursement in the basic benefit package means available to all, so we cannot restrict reimbursement to only patients that participate in the study” (*Netherlands policy maker 1*). Similarly, all new devices

used in the inpatient setting can be reimbursement through short-term ‘NUB payments’ in Germany regardless of additional evidence collection, so the “right incentives may be lacking to recruit patients” (*German policy maker 1*). In Germany, CED may therefore only be effective in the ambulatory care sector for devices used outside of the hospital.

There are also unique characteristics to the device sector that may discourage CED studies. For SMEs, who largely comprise the medical device market, there is “considerable pressure to get their devices to the market quickly to bring in revenue” and, given the shorter development timeframes for devices, avoid the “threat of fast followers that can get on the market before technology completes the study” (*UK policy maker 2*). In relation, there is an issue of a “free rider problem”...“if manufacturers think that data collected on a similar device can be used to obtain coverage for their device, then they will not undertake their own study” (*US policy analyst 4*).

Applying new evidence to coverage policy

Applying CED within the coverage process can create time pressures on designing, conducting, and analysing studies. In most cases, evidence must be generated in a limited time frame to inform a final coverage determination. Some informants maintained that the time frame for CED studies is generally too short, which prevents the collection of sufficient data. “Studies simply do not always generate needed data” (*Netherlands policy maker 1*). Even in the case of new CED schemes, there are “doubts that the data generated will be conclusive enough to effectively inform coverage policy” (*Germany policy maker 2*). For example, an RCT on negative pressure wound therapy was initiated a couple of years ago in Germany to gather further evidence on its performance in the community setting, with the goal to recruit all patients by the end of 2014. Given patient recruitment rates to date, one of the informants familiar with the study estimated that only around 50% of the total target number of patients will be recruited by that date. Conversely, some informants indicated that the time frame for CED studies “hinders their utility” (*US policy analyst 3*), as data collection takes too long to inform decision making in a timely manner.

The policy impact of several CED studies has been somewhat limited to date given that study timelines are extended due to the time required to approve studies and/or recruit patients. In some cases, studies do not start or are continuing to run more than five years after the initial CED decision. Some of the delay in achieving reasonable timelines is linked to difficulties in agreeing the terms of the CED scheme, securing ample funding for studies, and the lack of incentives for physicians to enrol patients. For instance, one informant mentioned internal problems with the perceived ownership of a particular registry. Other times, studies may not get underway due to political and practical issues associated with revising coverage. “Insurance companies have expressed some reservations about the conditional entry policy, because of concerns that it will be hard, if not impossible, to stop reimbursement once it is given, even if the resulting evidence demonstrates it is less effective than anticipated or in cases where the evidence is inconclusive” (*Netherlands policy maker 1*).

To that end, a key issue that hampers the application of evidence to coverage decisions once studies are completed is a lack of relevant or conclusive data. “Some studies ultimately did not provide the type of evidence Medicare needed to make informed decisions” (*US policy analyst 2*)...“when data are missing or resulting data is inconclusive, it is difficult to stop reimbursement” (*Netherlands policy maker 1*) and, “Then what? Accept outstanding uncertainty and make a decision whether to fund, acknowledge poor study design or inadequate patient recruitment, or continue with further evidence collection?” (*UK policy analyst 2*).

Ideally, the evidence generated by studies will be clear and conclusive. However, even in such cases, ensuring its use in coverage decisions can be problematic, especially if the data are negative. “If [the evidence] is clear, it is easy to make a decision, but then there are challenges with acceptance of the decision among clinicians and patients, if the decision is negative or calls for restrictions on access” (*Netherlands policy maker 1*)...“It is nearly always very hard to modify or remove coverage at a later date, unless there is a complete lack of effectiveness or important safety concerns” (*UK policy analyst 3*). Despite these political challenges, informants maintained that if the results of CED studies do not support use, the

technology must be withdrawn from coverage if the policy is to have impact and provide the necessary incentives for evidence generation....”If results are negative, the device has to be withdrawn from reimbursement, but this is not always the case and it hinders the impact of CED and stakeholder buy-in to support additional evidence generation. The policy needs to reward those that undertake studies and meet outcome expectations” (*Industry representative 5*).

Some of these issues played out in the CMS review of lung volume reduction surgery (LVRS). Rapid growth in the procedure volume of LVRS, despite limited evidence supporting its safety and effectiveness prompted CMS to suspend payments and co-sponsor a nationwide RCT (the National Emphysema Treatment Trial) to evaluate the procedure (Ramsey and Sullivan 2005). The subsequent trial findings showed no improvement in survival for surgical patients, but differential improvements for a percentage of surgical patients. Moreover, the subgroup analysis resulted in uncertain significance and the cost-effectiveness of LVRS compared with medical therapy was judged to be unfavourable. However, CMS’ resulting coverage policy was considered generous based on both the weight of the evidence and incremental benefit for LVRS patients show in the trial. Some analysts have questioned the politics involved in shaping the agency’s decision, suggesting that it was swayed by political, professional, advocacy, and other pressures (Gillick 2004).

Discussion

While CED is increasingly discussed in health policy and scientific circles as a potentially beneficial way to improve coverage policy, there has been limited empirical investigation of this approach. Therefore, this research fills an important gap, serving as the first study to examine the role and operation of CED policies in the context of medical devices. This work is especially important given that many international CED policies or programs are increasingly focusing on devices or procedures involving devices.

Countries in Europe and elsewhere have increasingly adopted CED policies to better inform coverage policies for new devices and other technologies. Overall, informants mainly agreed that CED holds potential to enhance coverage decisions and

strengthen existing evidence bases on the benefits and costs of new technologies, with resulting benefits to payers, manufacturers, hospitals, physicians, and patients. CED was deemed a particularly appropriate mechanism for devices, given their sometimes-underdeveloped evidence base and unique and diverse characteristics. However, use of the approach in practice is not without challenge, as previously discussed. There are a number of areas for improvement and possible strategies that could be considered to help address these issues.

First, there is a need to attain greater predictability associated with CED policies. Countries that have recently established CED programs, such as Germany and the Netherlands, have strived to make the process more standardised and transparent by clearly outlining eligibility criteria for CED and the initial procedures for consideration of applications or selected CED topics. However, after selecting technologies for CED and developing a call or request for research, almost all countries struggle to maintain a clear and predictable process. Therefore, formal and transparent procedures are needed to more effectively determine, execute, and communicate study design issues, the oversight and funding of studies, actions to address potential setbacks (e.g. slow or insufficient patient enrolment, insufficient or inconclusive data) and how generated evidence will be used to inform coverage decisions.

Such procedures may include requiring a series of pre-study meetings with involved parties to discuss key study design considerations upfront, including what endpoints will be collected, comparators used, criteria for study site selection, patient eligibility criteria, arrangements about access to data, opportunities for stakeholder input, and what constitutes an acceptable outcome(s) and how that will be assessed over what time period. Mechanisms to periodically provide publicly-available updates on the status of studies and report the results would also be beneficial. For example, in the US, if the CED scheme involves an RCT, information on the status of the study can be found on ClinicalTrials.gov, but providing current updates where the scheme is reported on the CMS website would be helpful, as well as reporting on non-RCT studies. The ISPOR Task Force on Performance-Based Risk-Sharing Agreements argues that CED and similar schemes have “public good aspects”, which should be considered. In part, this means that public authorities that negotiate and

fund CED should make the results of the results public, where possible (Garrison et al. 2013).

In addition, payers responsible for CED policies could produce and disseminate clear guidelines outlining the overall process and various stages that manufacturers, health care professionals, patients, and others, such as national research organisations or independent academic groups, can expect when a CED scheme is undertaken. The aforementioned actions could go some way in enhancing commitment from industry to engage in CED studies, ensuring reasonable timelines are achieved, and improving the quality of the data collected and, in turn, the ultimate coverage decision.

Second, policy makers, manufacturers, and the research community need to be better aligned and collaborate to ensure that CED studies most effectively address the outstanding uncertainties in the most efficient and timely way possible. Results from collaborative CED studies can be viewed as more credible and neutral, than those from studies undertaken by a single payer or by industry alone. Moreover, although payers may be well suited to identify the need for evidence, there are other critical assessments (e.g. type of research and its priority) that are not necessary ones for which they have particular expertise. Informed judgements and better decisions might be possible through greater involvement of the research community. The case of the ICD CED scheme in the US is a good example of the potential value of bringing together these various stakeholder groups. The ICD registry would not have been possible without the cooperation and financial support of the ICD manufacturers or the electrophysiology community, which supported the ICD registry primarily as means to monitor and potentially minimise procedure-related complications associated with ICD implantation. Private payers were also brought into the process and agreed to help fund the ICD registry and contribute ideas for what data elements would be useful for future coverage decisions.

To date, CED initiatives have largely been collaborative, although there is room for improvement. For example, Germany's new CED policy is offering opportunities for greater collaboration with and within the medical device industry. Manufacturers will be encouraged to exchange ideas about study design with the

GBA and to collaborate together on joint applications. This is particularly beneficial given that CED studies only apply to procedures, where a procedure could involve various types of the same device. Another benefit across all countries entails opportunities for manufacturers to explain their technology and any particular issues that may impact the design of a study to payers and others, especially in systems where manufacturers are not involved in the actual conduct of the study. Besides enhancing the involvement of manufacturers, more opportunities for co-sponsored studies should be established. For example, payers could support any patient care costs, while pooled funds from product developers or public research grants could support the research costs. In the US, the CMS could work more closely with other federal agencies, such as the FDA, NIH, and AHRQ, to identify topics for CED and oversee and fund studies. Moreover, PCORI could provide support for ongoing registry or clinical trial systems. In all countries, the involvement of clinicians or medical associations may prove particularly helpful, given their early involvement in device development and acquired early knowledge of particular technologies. As part of any discussion among the interested parties, it is important to gain a shared understanding of the reasons for the particular CED scheme and the ways in which the information gathered would reduce the uncertainty around the coverage decision.

Third, better incentives are needed to encourage physicians to recruit patients and/or collect data associated with CED studies. One possible strategy is to more adequately compensate physicians for referring patients to studies, especially clinical trials. Alternatively, it may be helpful to exercise some kind of negative sanction against physicians that do not comply with patient recruitment or data collection requirements. Making it easier for community-based physicians to participate actively in clinical trials could also have a positive effect on patient recruitment in CED studies, as well as improve the engagement of the community in important research and increase the chances that physicians will change their practice patterns based on the research results they were involved in generating. In addition, to encourage physician participation, study questions and protocols should be designed in the context of clinical practice, meaning that the procedures required by the study protocol should be easily incorporated into practice. Finally, it could be helpful to utilise non-traditional patient enrolment strategies alongside more traditional approaches, such as using physician referrals. Such tactics include using social media

networks, such as Facebook, Twitter, and YouTube, online data mining, and electronic medical record monitoring and analysis.

Fourth, given the time and resources required for CED schemes, it is important to capitalise on existing data collection sources and networks to monitor technology use and patient outcomes after the initial coverage decision has been made. In particular, CED efforts should align with existing mechanisms to expand electronically-available health data, including longitudinal patient registries, electronic health records, and, in the US, claims data collection and analysis. In Europe, several jurisdictions, such as Germany and the UK, have introduced patient registries, particularly in orthopaedics and cardiology, to collect post-market data, which are usually collaboratively supported by medical associations, academic centers, and national research organisations. Again, in the US, such actions would include creating greater synergies between the CMS CED program and current investments in CER. Together, these efforts would help ensure that the necessary data is generated to support CED schemes as effectively as possible and, ideally, that better evidence exists to make informed coverage decisions in the first instance.

Fifth, and in relation to the previous point, one strategy to improve the effectiveness of CED or perhaps reduce the need for it is to address some of the challenges related to pre-market regulation of devices. In particular, as previously discussed, data on the effectiveness of devices (particularly with regards to high-risk Class III devices) is not often available or is insufficient at the time of coverage determination. These are issues that are largely attributable to the fact that such data are not commonly required for a device to achieve market authorisation. While it is outside the scope of this paper to discuss potential regulatory reforms to address this issue⁷⁹, one approach to consider is for payers and regulators to provide scientific advice to manufacturers to better harmonize evidence requirements. In the US, for example, the FDA and CMS have initiated a voluntary, two-year “parallel review” program for devices, which entails a partial alignment of their respective review processes for regulatory approval and coverage, respectively (Messner and Tunis 2012).

⁷⁹ See Sorenson C, Drummond M. (2014). Improving medical device regulation: The United States and Europe in perspective. *Milbank Quarterly* 92(1): 112-148.

Sixth, several of the informants noted the need for greater international collaboration between those involved in national CED schemes. At present, devices are being studied in multiple countries simultaneously, which duplicates efforts and consumes resources that could otherwise be dedicated to applying CED to a greater number of technologies or towards other objectives. Increased collaboration between countries would also help lower-resourced countries to conduct CED in a more regular and formal manner. The EUnetHTA has played some role in fostering greater collaboration between European countries and information exchange about what technologies have been studied and the associated research and policy outcomes. For example, the EUnetHTA Planning and Ongoing database allows HTA agencies to share information with each other on planned and ongoing projects conducted at the individual agency. In addition, the Evidence Database on New Technologies allows sharing and collection of information on coverage and reimbursement and assessment status of promising technologies and on additional studies requested or recommended further to an assessment.

Finally, it appears that almost all countries struggle with ensuring that evidence collected in CED studies is subsequently used to inform coverage policy. A range of factors impact the resulting use of evidence, including the quality and certainty of the data generated; alignment between the timing of data collection and the decision needs of policy makers, physicians, and patients; and, political acceptability, particularly if the evidence points to removing coverage. Some of the aforementioned actions would help towards addressing some of these issues. However, CED is only effective if it is used to inform decision making and, therefore, a coverage decision must be made reflective of the data generated and in alignment with any previously agreed commitments between all involved parties. This is a necessary step to foster trust and commitment in, and, ultimately, the utility of the CED approach.

In closing, it is worth noting the limitations of this study. The opinions and experiences of the experts we interviewed may not be reflective of all those involved in CED schemes internationally. However, the experts involved in the study are leading internationally-recognised experts in the areas of HTA, CER, evidence-based

policy making, and coverage and reimbursement policy. Those from payer agencies and industry are senior leaders within their respective organisations. Moreover, our informants constitute a variety of key stakeholder groups involved in CED to ensure that a range of perspectives and experiences was captured. In addition, the opinions of the experts interviewed were complemented by, and verified against, information gathered from our literature review.

Conclusions

CED offers potential to provide patients with access to potentially beneficial technology, while duly enhancing the existing evidence base on the technology. As an approach, it also has the benefit of improving evidence-based coverage policy. However, as highlighted throughout this paper, employing CED in practice comes with various challenges, from determining eligibility criteria to designing and funding studies to applying the new evidence to coverage decisions. Some of these challenges have presumably stymied the use of CED, as evidenced by the limited use of this approach to date in most countries. The strategies discussed herein may help towards identifying optimal use and supporting improvements in the operation, and implementation of CED. Given the dynamic nature of medical technologies and CED itself, the approach should be evaluated on both a short-term and long-term basis, in terms of its impact on static and dynamic efficiency.

Study 7: Valuing end of life care: The case of advanced cancer drugs in the United States

Introduction

New cancer therapies offer the hope of improved prognosis to patients with life threatening disease. Over the past five to ten years, a number of specialty treatments in particular have entered clinical practice in the US to provide better systemic therapy for advanced cancers that respond to few therapeutic alternatives. To date, however, such advances have been only modestly effective in extending life (Schnipper et al. 2010; Schrag 2004).

Alongside the optimism generated by new advanced cancer drugs comes difficult trade-offs. The prices of many of these therapies exceed \$25,000 a year and result in benefits measured in months (Fojo and Grady 2009; Meropol and Schulman, 2009). Consequently, patients are often faced with exorbitant costs and physicians are increasingly placed in the undesirable position of having to help patients decide whether the potential benefits warrant the financial strain that these medications may generate. Concern over the high costs and relative value of new cancer drugs are not only confined to physician offices, but also in discussions at influential meetings of the American Society for Clinical Oncology (ASCO) (Meropol et al. 2009) and on the pages of medical journals and the popular press. Moreover, such issues are being considered against the backdrop of recent health care reforms – from more general discourse on costs, access to care and obtaining better value from the stratospheric spending levels on health care to more specific proposals around the use of CER and payment reform. The dramatic trade-offs proffered by these drugs extends beyond the national arena; indeed, their adoption has become a touchstone for broader health policy debates elsewhere, most notably in the UK (Chalkidou 2012; Department of Health 2010; Faden et al. 2009; Jeffreys, 2007).

With increased attention worldwide placed on the cost of cancer care, the time is ripe to explore how cancer drugs are currently being valued in the US context. This article aims to address this issue, particularly in the case of end-of-life care, and

is structured as follows. First, the article provides some context around these new therapies, followed by a discussion on costs and patient access. The article subsequently turns to outlining the current evidence on the value for money of advanced cancer drugs and then examines some of the technical, political and social challenges in assessing their value and considering such evidence in decision making. The article concludes with some preliminary thoughts on how to obtain greater value in advanced cancer care.

Waging the war on cancer

“.....the time has come in America when the same kind of concentrated effort that split the atom and took man to the moon should be turned toward conquering this dreaded disease. Let us make a total national commitment to achieve this goal.”

President Richard Nixon, January 1971, State of the Union Address

Eleven months later, President Nixon formalised the war on cancer with the passage of the National Cancer Act, an act designed to promote the discovery of new treatments for cancer and to encourage early detection and prevention of disease. More than three decades later, cancer remains an important cause of mortality and morbidity in the US. In 2007, cancer represented the second leading cause of death, accounting for 23.2% of all deaths, or roughly 563,000 people (Centers for Disease Control and Prevention 2010). Today, 1.5 million new cases of cancer are diagnosed in the US annually (American Cancer Society 2010). However, the last three decades have witnessed important strides in understanding, preventing and treating the disease. Scientists have identified some of the genes responsible for cancer, discovered new chemotherapeutic and biologic approaches to treating the disease and developed new imaging techniques to detect cancer earlier. Accompanying these scientific advances have been equally important public health campaigns and educational programmes, which have led to improved screening rates and lifestyle changes (e.g. reduced smoking, increased physical activity) that help prevent the occurrence and reoccurrence of disease. Consequently, median survival rates have improved for many cancers, particularly breast cancer, colon cancer, prostate cancer and non-Hodgkin's lymphoma (American Cancer Society 2010).

Despite these advances, however, there is significant debate over whether the war is truly being ‘won’ and, if so, at what cost. Criticisms generally centre on two closely related issues. First, there are concerns that medical progress in cancer care have predominately focused on the development of expensive treatments that only marginally prolong life and may reduce its quality (Epstein 2005; Faguet 2005). As intimated previously, many commentators have questioned the cost-effectiveness of recent cancer treatments and whether the high cost of such therapies is therefore warranted (Berenson 2005; Faguet 2005; Kolata and Pollack 2008; Malin 2010; Shih and Halpern 2008). Critics often point to the fact that the US spends twice as much as any other country for the same overall survival results (Meropol and Schulman 2009) as evidence that value in advanced cancer care is suboptimal. Second, some argue that too much effort has been placed on developing such costly treatments, at the expense of prevention and early detection (Faguet 2005). The following sections critically explore the first issue in further detail.

The price tag on progress: Costs of and access to new cancer drugs

Expenditures on cancer care have nearly doubled over the last 20 years, in part as a result of the prices and rapid uptake of new agents and other technologies, including advances in imaging and therapeutic radiology (Meropol et al. 2009; Tangka et al. 2010). In particular, spending on cancer drugs has risen 15% annually in recent years, with many approved medications carrying costs of \$5,000-\$10,000 or more per month of treatment (Gatya and Longwell 2008; Schickedanz, 2010). Sales of cancer drugs are now second only to those of drugs for heart disease and estimates suggest that more than 40% of Medicare drug spending is for oncology agents (McNeil 2007; Smith and Hillner 2011). To be sure, scientific advances in oncology have not come cheap.

The growing economic burden associated with such treatments begs the question of why costs are so high. First, there are factors related to their supply and demand that contribute to high drug prices. Most new cancer drugs are biologics, which are more costly to produce than traditional medications and administered via infusion or direct injection, both of which involve substantial costs. The rising costs of production is partly due to the escalating expense and inefficiencies of clinical

trials and the time and resources required to meet the evidentiary requirements of national drug regulatory authorities (Rawlins and Chalkidou 2011). In addition, the majority of therapies are patented and under other market exclusivities that limit competition in order to preserve incentives for research and development (R&D), thereby reducing competitive (downward) pressures on prices (Danzon and Taylor 2010). Furthermore, as many of these drugs treat advanced disease that respond to few therapeutic alternatives, they are able to command premium prices without significantly deterring provider and patient demand (Goldman et al. 2010). Neumann et al. (2010) and Nadler et al. (2006) found that 67% and 78% of oncologists, respectively, endorsed the notion that all patients should have access to available and effective care regardless of costs. For many terminally ill patients, even when prices are high, demand for cancer drugs has been found to be largely inelastic, as they understandably place a high value on any additional life and therefore willing to pay more for treatment (Goldman et al. 2010).

Second, existing health insurance distortions and payment policies contribute to high costs. In principle, insurance creates a gap between the out-of-pocket prices that consumers pay and those received by manufacturers, with the difference paid by insurance. In cases where insurance coverage is insufficient to cover costs, many patients have supplementary insurance to cover the remainder or they forgo treatment if they cannot afford the co-insurance, and manufacturers often offer discounted drugs or assistance programmes for some patients without the ability to pay (Danzon and Taylor 2010). Consequently, patients and physicians have had limited incentives to demand cheaper drugs, thereby giving manufacturers little reason to constrain prices, although this may change due to the rising cost of care and growing levels of patient cost-sharing, as discussed further below. In the case of Medicare, current payment policies for cancer drugs used under both Parts B and D benefits further encourage rising costs. With regard to the former, existing policies effectively reduce physicians' margins⁸⁰, which encourages manufacturers to charge high prices,

⁸⁰ Most cancer drugs are dispensed in physician offices and therefore covered under Part B or by a private insurer's medical benefit. Prior to 2005, they were paid at average wholesale price (list price), but are now paid at average selling price (ASP) plus 6%. Historically, physician reimbursement for oncology treatments has been highly lucrative, based on the relative margin between acquisition costs and reimbursement rates (the 'buy and bill' approach).

especially at initial launch, to offer larger margins to physicians and indirectly influence use of costlier, more aggressive therapies (Danzon and Taylor 2010; Jacobson et al. 2006; Medpac 2010). In contrast, reimbursement for cognitive services (e.g. end-of-life counselling) is dismally low (Smith and Hillner 2010). While there is little to no evidence that physicians base treatment decisions solely on potential profit, when deciding between two equally efficacious treatments, oncologists tend to choose the more expensive therapy (Jacobson et al. 2006; Smith and Hillner 2010).

For drugs under Part D, pharmacy benefit managers contracted by Medicare have limited ability to negotiate price discounts for preferred formulary placement with certain classes of drugs, including those for cancer (Danzon and Taylor 2010). In fact, as of 2010, contracting private plans are required to include all drugs for conditions that are major or life-threatening (Bach 2009), regardless of cost or value for money. To that end, Medicare's overall standard for coverage is whether the drug is 'reasonable and necessary' for the diagnosis or treatment of disease; it is not permitted to refuse coverage of a drug based on grounds of either cost or cost-effectiveness. A further source of upward pressure on costs is the extensive use of and payment for 'off-label' cancer drugs and treatments within Medicare⁸¹. While off-label use in cancer care is commonplace, with estimates indicating that 50–75% of drug or biologic therapy used to treat cancer is off-label (Soares 2005), use is higher in advanced cases when patients are no longer experiencing benefit from standard approved treatments, and increasingly involves expensive therapies (Sullivan et al. 2011). Some off-label use is supported by clinical data, although in many cases there is little supporting evidence of benefit, with accumulating evidence of harm (Giezen et al. 2008; Hecht et al. 2009; Tol et al. 2009). Nevertheless, off-label use of cancer therapies has endorsement from the NCI and, in some circumstances, the FDA. Moreover, the CMS must pay for off-label use if available medical compendia – reference guides Medicare relies on to determine which off-

⁸¹ Off-label use entails any use of a drug different from that described in the FDA-approved drug label. In oncology, off-label use includes uses for a different cancer or at a different time in the course of the disease, or in a dose or schedule different from that in the approved label.

label uses of cancer drugs to cover – recommend use, even if such recommendations are based on scant evidence of effectiveness and safety⁸².

Given that private plans generally follow Medicare’s policies and coverage and reimbursement decisions, the aforementioned payment rules have a ripple effect across the whole of the health system, affecting physician treatment choices and distorting pricing on value and broader cost containment.

Thirdly and finally, the high costs associated with advanced cancer care are also attributable to overutilisation of treatment or the utilisation of futile care. Several studies suggest that a substantial portion of the total cost of cancer care is for treatment delivered in the last months, weeks, or days of life, and that much of this care is of little to no therapeutic benefit and potentially inconsistent with patients’ wishes (Earle et al. 2008; Lubitz and Riley 1993; Mack et al. 2010; Zhang et al. 2009). Existing evidence indicates that patient preferences for end-of-life cancer care vary widely (i.e. some patients want aggressive treatment until the very last day, others want home or hospice care with minimal medical intervention) and that existing national variations in care⁸³ largely do not reflect such preferences or differences in need (Barnato et al. 2007; Earle et al. 2008; Goodman 2011; Goodman et al. 2010, 2011). Rather, regional differences in end-of-life care intensity are often guided by ‘supply-sensitive care’, resulting from uncertainty about how best to treat advanced cancer patients, practice styles of different health systems and physicians, and the tendency of physicians to use the medical resources available to them, especially when it involves lengthening life⁸⁴ (Goodman et al. 2011; Hollingsworth et al. 2010).

⁸² Compendiums entail a listing of drugs, their clinical properties and recommended uses, and are put together by non-governmental bodies (e.g. American Hospital Formulary Service Drug Information). If only one compendium recommends use, Medicare must provide coverage. Based on a systematic review of available compendium, Abernethy et al. (2009) noted variability in the consistency and transparency of the policies applied in delineating off-label cancer indications and the approach to the review, rating and updating of evidence was often neither systematic nor comprehensive.

⁸³ The use of chemotherapy and life-sustaining treatment at the end of life varies considerably. Goodman et al. (2010) found that the percentage of cancer patients receiving chemotherapy during the last two weeks of life varied fourfold among hospital referral regions and the percentage receiving life- sustaining treatment during the last month of life varied by a factor of more than six. Similar patterns were found for intensive care and the use of hospice care (Goodman et al., 2010, 2011).

⁸⁴ This may be due in part to the basic tenets of practicing medicine – that is, to sustain life – as well as the changing norms of modern medicine, which entails using more technological approaches to care over other means.

The growing price tag associated with these treatments is increasingly impacting patients and providers alike. Even Medicare beneficiaries and those with private health insurance can be faced with considerable out-of-pocket costs. Goldman et al. (2010) reported that although median patient out-of-pocket costs for new oncologic biologic therapies was only 1-2% of the total cost, the mean was 4-13% at the 90th percentile, and patient share of total cost ranged from 13% for imatinib to 15% for rituximab. For Medicare beneficiaries in particular, Faden et al. (2009) found that patient costs can total between \$2,900 and \$6,000 for three months of treatment to \$7,400 and \$24,000 for an annual course of therapy. Such co-pays are particularly a concern for those with low or fixed incomes, who spend about a quarter of their annual income on such expenses (Berenson 2005; Kim 2007; Langa et al. 2004). For uninsured and underinsured Americans, out-of-pocket costs can easily be four times higher; depending on the agent, annual therapy, for example, can soar to more than \$100,000 (Faden et al. 2009). Given that patients are likely to be taking multiple drugs, patient expenditures are presumably higher.

When confronted with high out-of-pocket expenses, some patients, particularly those without the ability to pay, may forgo or discontinue treatment (Szabo 2008). Even those patients with some capacity to finance treatment may not undertake all recommended care in fear they will burden their families with unmanageable debt (Berenson 2005; Kim 2007). Among those declaring bankruptcy for medical reasons, cancer is frequently the condition that precipitates the financial crisis (Himmelstein et al. 2009). Therefore, while increased patient cost-sharing may serve to control overall costs, there is evidence that it is associated with worse outcomes in the sickest and poorest patients, perhaps by causing lower use of necessary services (Hsu et al. 2006). Growing patient cost burdens is also affecting physician treatment patterns. In a survey of oncologists (Neumann et al. 2010), 56% and 84% reported that the cost of new cancer drugs and patient's out-of-pocket costs, respectively, currently influence which cancer treatment they recommend to their patients.

Evidence on the value for money of end of life cancer drug treatments

Given the substantial costs associated with advanced cancer drugs, their value for money is increasingly being questioned by patients, clinicians, payers and policy makers, with many observers arguing that these drugs, often used in the last months of life, are not cost-effective, in the sense of extending QALYs at a reasonable price. Indeed, these treatments frequently have cost-effectiveness ratios that exceed (or well exceed) \$100,000 per QALY, which is higher than the \$50,000-\$100,000 per QALY threshold often suggested as denoting good value for money (Neumann et al. 2000). Suggestions have been made for the threshold to be raised to take into account societal preferences and medical inflation (Ubel et al. 2003), and that cancer drugs (that increase life expectancy in the final month of life) in particular should be regarded differently than treatments that lower mortality risks only incrementally throughout an individual's entire adult life (Becker et al. 2007). As discussed in Chalkidou's (2012) article in this issue, the national HTA body in the UK, NICE, recently extended its cost-effectiveness threshold for drugs aimed at end of life care under certain circumstances. The new policy was, in part, guided by societal preferences for the lives of terminally ill, mostly cancer, patients to be valued more than those suffering from other, potentially curable, chronic, or acute conditions.

However, there is limited understanding of how these treatments are valued by different stakeholders in the US, although some evidence is starting to emerge. In a recent survey of oncologists on the value of cancer drugs, Neumann et al. (2010) found that they had an average implied cost-effectiveness threshold of roughly \$300,000 per QALY. Interestingly, when later asked to generally define reasonable value for money, the majority (49%) indicated \$50,000-\$100,000 per QALY, with approximately 20% deeming a higher range of \$100,000-\$150,000 per QALY as good value. A similar study (Nadler et al. 2006) demonstrated that 62% of oncologists believed that a life expectancy gain of at least two to four months justified the use of a hypothetical agent with a cost of \$70,000 per year above the standard of care and another 20% believed four to six months would justify this cost.

Goldman et al. (2010) estimated patients' willingness to pay for cancer drugs⁸⁵ to be equal to four times the total cost paid by the patient and his or her insurer. Annual net benefit ratios were slightly higher for those aged 65 years or older and female patients; no differences in willingness to pay were found by income, although there were acknowledged limitations in how income was imputed. Research by Becker et al. (2007) and Selvin et al. (2010) highlight that people who face imminent death, as in the case of advanced cancer patients, may place a much higher value on life-extending treatments than people for whom death is a remote risk. This may result in a willingness to spend any remaining wealth to live a few weeks longer, because the value of their wealth may lose much of its value at death.

Despite this initial evidence suggesting high valuations for advanced cancer drugs, there are qualifications to making deductions about their value for money. Importantly, oncologists and terminally ill patients may have different perceptions regarding economic value and benefit than other physicians, patients and broader society. For example, oncologists often practice in a context where grave situations in which treatments offer modest benefits are the norm and this arguably influences their perceptions of value (Nadler et al. 2006). Smith and Hillner (2010) suggest this may indirectly cause oncologists to overvalue (and overutilise) costlier drugs, as they expect sufficient financial reward in return for routinely dealing with complicated treatments and the "fractured dreams and extreme suffering" of patients.

Furthermore, Nadler et al. (2006) found that physician perceptions of economic value do not correlate highly with their own practice patterns. The majority (78%) of oncologists surveyed believed that patients should receive 'effective' treatment regardless of cost, but when queried whether a therapy that has been shown effective offered good value, most were unsure or disagreed, which suggests that they offer drug therapy to patients regardless of perceptions of value for money as long as the treatment is considered effective. Recent research by Colosia and colleagues (2011) highlights that oncologists' perceptions of quality cancer care includes effectiveness, but not costs. In addition, physicians often either under- or over-estimate survival benefits and costs, indicating that some oncologists' knowledge on the matter is lacking or that they have exceedingly high cost-effectiveness thresholds. How these

⁸⁵ Goldman and colleagues looked at Avastin, Herceptin, Rituxan, Tarceva and Gleevec in particular.

findings comport with evidence that oncologists are increasingly considering cost in treatment decisions is worth further exploration.

Similar issues hold for advanced cancer patients, but are likely heightened given their lack of medical expertise, variable understanding of the risks, benefits, and costs involved⁸⁶, and the distress of facing death, which as stated earlier, may result in higher valuations of these therapies than would be elicited by healthy people or patients with acute or chronic conditions.

In addition, given increasingly limited resources to fund all health care, the value of cancer drugs should be considered from a broader societal perspective, which has yet to be explored in the American context, to ascertain how public valuations of these treatments might differ from cancer patients and their physicians. Determinations of their social value should be gauged against that of other interventions or policy priorities (inside and outside of health care), many of whom may provide better value for money.

The road to value-based funding for and access to cancer drugs: Technical, political and social challenges

It is clear that use of these drugs at the end of life often entails a complex cost-benefit trade-off for physicians, patients, payers, and society alike, which will only get more challenging in the future with greater R&D activity in this area, increasingly expensive therapies⁸⁷ (Schickedanz 2010) and tighter health care budgets. In attempts to address the “value crisis in oncology” (Ramsey and Schickedanz 2010), there have been recent developments toward better understanding the value of cancer drugs and measuring the value of cancer drugs and advanced cancer care more generally.

⁸⁶ Patients’ ability to understand probabilities of survival and adverse events are influenced by their health numeracy skills (Ancker and Kaufman 2007) and patients tend to update their perceptions of the benefits and harms as well as the level of uncertainty based on their individual clinical experiences and subjective assessment of treatments and resultant outcomes. Harrington and Smith (2008) found that one third of patients persist in believing they can be cured even when conversations and evidence stating otherwise are documented.

⁸⁷ Currently, there are over 100 new cancer drugs in randomized phase III trials, all with a price tag of thousands of dollar per month.

In particular, in the last few years considerable attention and investment has been put towards CER. The ACA dedicated \$1.1 billion to fund CER⁸⁸ and establish the public-private PCORI to oversee the research. The aim of CER is to provide evidence of which interventions work in practice to allow physicians and patients, among others, to make better health care decisions. Cancer was one of the key priorities identified in early CER priority-setting processes (Institute of Medicine 2009) and recent research suggests that there is demand for CER on cancer drugs. In particular, 79% of oncologists surveyed wanted more government research on their comparative effectiveness and 80% favoured greater use of cost-effectiveness data in coverage and payment decisions for these treatments (Neumann et al. 2010).

While CER is arguably needed in the context of cancer drugs and other available interventions – less than half of medical care in the US is based on or supported by evidence about its effectiveness (IOM 2007) – there are aspects of cancer drugs that introduce technical challenges to measuring their value (Mullins et al. 2010). Due to the incentives inherent in the clinical research enterprise, RCTs are designed for specific purposes of regulatory approval and maximum market penetration, not to reflect the complexities of real-world patient care, where patients may have significant co-morbidities and other clinically mitigating factors, and they often do not collect economic information. Moreover, trials are rarely designed to compare a new drug against existing treatment alternatives to demonstrate relative therapeutic advantage (Smith and Hillner 2010).

Although these are not challenges unique to cancer drugs per se, for a variety of reasons these issues are exacerbated in the area of oncology. In cancer care, for example, clinical uncertainty in measuring comparative effectiveness leads to economic uncertainty, as for many cancers, there is no dominant treatment alternative and thus each attempt to treat incurs a separate cost, especially toward the end-of-life when additional, more aggressive lines of therapy are employed (Virgo et al. 1995). It is also difficult to accurately capture patient preferences and assessments of impact on quality of life when, in advanced cancer, there are notable uncertainties about survival, side effects of drugs, and personalised responses to care that

⁸⁸ The ACA defines CER as “research evaluating and comparing health outcomes and the clinical effectiveness, risks, and benefits of two or more medical treatments, services, and items”.

frequently change along the course of treatment. Along these lines, existing preference-weighted quality of life measures (often used in QALY determinations) may not be sensitive enough to pick up changes in the health status of cancer patients or reflect preferences driven by specific characteristics or severity of the disease (Garau et al. 2010). In addition, the ubiquitous use of oncology drugs off-label makes it difficult to assess their value, as there is often insufficient or poor quality evidence from clinical trials to determine benefit, as previously discussed (Mullins et al. 2010).

There are also political challenges to measuring value and, in particular, to applying resulting evidence in decision making on the availability and use of cancer drugs. The emotional debates preceding passage of the health reform legislation illustrated strong disagreements about what should constitute ‘value’ (i.e. whether both clinical effectiveness and costs (or other factors) should be considered), who should evaluate value, and what role for CER in decision making (Garber and Sox 2010). Some legislators opposed using CER evidence for federal decisions about health insurance coverage, a stance supported by intense lobbying from various interest groups and dramatised with references to ‘death panels’, the dreaded R-word (rationing), and the heartless British NHS who leaves grandma to die to save money. These debates highlighted the political risk of giving PCORI even the resemblance of authority over health care decisions and, consequently, the ACA prohibited the Institute from assuming the responsibility or authority to measure value and limited the ways in which CER can be used by Medicare (Garber and Sox 2010). Moreover, it barred Medicare from using cost-effectiveness as a factor in making coverage and reimbursement decisions⁸⁹. Such provisions were largely supported by the general public (Gerber et al. 2010).

Research specific to advanced cancer drugs suggests oncologists concur; very few desire explicit resource allocation with government assuming a role in determining what care constitutes good value for money (Neumann et al. 2010). While current provisions rendered CER more politically palatable and fostered

⁸⁹ The bill specifically prohibits any cost-effectiveness analysis that would use any adjusted life years factor that would place lower value on the life of elderly, disabled, or terminally ill individuals compared to younger and healthier persons, as well as the use of a strict cost-effectiveness threshold.

stakeholder acceptance of PCORI, there is risk that without some policy ‘teeth’, such research will have limited impact on clinical practice and costs.

Given the existing scenario where CER is restricted to more decentralised and implicit use by physicians, patients, and other actors in the health system, there has been a surprising lack of discussion of how CER evidence on the value of treatments will get from ‘bench to bedside’. Simply making such evidence available will not likely be sufficient to ensure its consideration in patient care, including end-of-life treatment – that is, evidence is only as good as what you do with it (Freemantle and Drummond 2010). First, in Neumann et al.’s study (2010) of oncologists, physicians reported insufficient time to consider available evidence, which may be heightened in end-of-life care when time to treat may be limited (Sullivan et al. 2011). Moreover, only 42% of oncologists said they felt well prepared to interpret and use such information in their treatment decisions. Second, even if one assumes that physicians and patients will have access to CER information on cancer drugs and use it in their decisions, there is evidence suggesting that the cost issue will not be adequately addressed. The most obvious reason is that as currently fashioned, CER evidence may not include the costs of therapy. Furthermore, patients are often reluctant to raise concerns regarding costs because they feel awkward addressing the topic and worry about how such discussions might affect quality of care (Alexander et al. 2003; Kim 2007). Oncologists feel similarly uneasy, which stymies open discussion of costs with patients (Neumann et al. 2010; Schrag and Hanger 2007). This discomfort may be more pronounced in the context of end-of-life treatment decisions, when it could feel that a life is being priced (Schickedanz 2010), especially if costs have not entered into patient-physician discussions previously.

In addition, the existence of a “cancer taboo” may pose additional political and social barriers to value-based decision making on cancer drugs. Cancer holds a privileged position - the “emperor of all maladies”,⁹⁰ - in US health care, as evidenced by significant public and private research investment in understanding and treating the disease, a powerful industry, the multitude of influential interest groups

⁹⁰ See S. Mukherjee (2010), *The Emperor of All Maladies: A Biography of Cancer*. This may be influenced by the research imperative that is pervasive in US health care (Callahan 2009). Considerable investments are made in medical research, particularly with regards to cancer, and researchers, funders, advocacy groups, among others, all have a stake in heralding the success of the research. Moreover, such investments uphold expressed beliefs in the incremental nature of medical progress and that small differences lead to large net gains in public health.

and national associations with considerable political and financial capital, and, of course, the human element that almost every citizen knows of someone who has suffered from or died from cancer. One only needs to envision the bevy of organised sporting events in support of various types of cancer and the ubiquitous coloured rubber bracelets that have now become a global phenomenon to understand the extent of public support for ‘fighting the war’. In the context where available CER evidence demonstrates a lack of benefit against the costs, such influence, coupled with national sensitivities toward rationing, will render saying ‘no’ extremely difficult. Inevitably, payers will fear negative publicity and potential lawsuits, policy makers will want to avoid ‘disease politics’ and the entire CER enterprise could be placed under threat.

While not specific to advanced cancer drugs, tensions between evidence, access to treatment, and political and social preferences played out in the recent breast cancer screening guidelines. The guidelines, if followed, would reduce the use of mammography among women ages 40 to 49; substantial opposition ensued, which led to a provision in the health reform legislation requiring the government to ignore the recommendations (Gusmano and Gray 2010). The experience of the UK, as discussed by Chalkidou (2012), further underscores the difficulty in balancing such tensions. Even in a long-standing and centralised system of CER via NICE, where the use of evidence to support access decisions is well known and formally integrated into policy, the Institute has faced significant negative media attention and public backlash to ‘no’ decisions for advanced cancer drugs, which arguably contributed to its change in policy for end-of-life treatments, the promulgation of a dedicated cancer fund, and discussion of potential removal of the Institute’s decision-making authority among government officials.

Conclusions

This paper illustrates some of the key issues present in valuing end-of-life care in the US in the case of advanced cancer drugs, from the difficult trade-offs between their marginal health benefits and high costs to the challenges in assessing their value and applying such evidence in decision making. While policy makers, payers, physicians and patients will continue to face such issues into the foreseeable future,

there are a number of initial steps that could be pursued to address existing evidence gaps and some of the outlined challenges.

First, as intimated earlier, additional research is needed to determine how different stakeholders value these treatments. While evidence specific to oncologists has recently emerged, there is a lack of understanding around the preferences of policy makers, payers, patients and the general public regarding the use of these drugs and within the context of end-of-life care in particular. This would be a first step in better understanding perceptions of value, how to effectively define and measure it, and, more broadly, how best to manage the different needs and expectations of these various groups.

Second, although cancer drugs introduce unique technical challenges to assessing their value, there has been movement to address some of the limitations of RCTs through the use of pragmatic clinical trials⁹¹ and CED, which has been used by Medicare since 2005 to provide conditional coverage for promising medical technologies while additional clinical evidence is generated to reduce the uncertainty and evidence gap in valuing treatments (Cohen and Looney 2010; Mullins et al., 2010). Additionally, better incentives are needed to address the poor evidence base with regards to off-label use, which could involve a greater role for Medicare to strengthen evidence compendia and for developing methodological standards for evaluating off-label indications (Sox 2009). Other strategies to build the evidence base in oncology, particularly with regard to better understanding what (and how) cancer care is delivered to patients and its consequences, include use of outcomes databases and developing improved metrics to measure quality of care. These approaches could be potentially incorporated into existing physician payment schemes to incentivise development of evidence about what works and for whom in advanced cancer care. It could also help guide and encourage appropriate care at the end-of-life, with the aim to protect against over- and under-utilisation of care. UnitedHealthcare recently introduced a pilot programme incorporating such

⁹¹ Pragmatic trials or prospective controlled studies are specifically designed to be informative for post-regulatory decision makers through selection of more relevant comparators, better generalizability to diverse patient populations, and inclusion of relevant outcome measures of interest to payers, clinicians, and patients (e.g. quality of life) (Mullins et al. 2010).

strategies, where oncologists can earn incentives for delivering quality care that results in controlling costs as well as for developing outcomes data (Burns 2011).

Third, although CER is focused on comparative effectiveness, it does not preclude consideration of costs. Garber and Sox (2010) argue that any CER research supported by PCORI and other sponsors should include data on use and costs to ensure that the most accurate and comprehensive information is available to decision makers. For example, CER-related legislation limits the use of cost-effectiveness analysis by certain federal programs, but cost data could still prove valuable to Medicare and other stakeholders. To the latter, private entities (e.g. private insurers, hospitals, physician groups) and other government programs could consider costs and perform or apply cost-effectiveness analysis based on comparative effectiveness results in their decisions. Insurers, in particular, could use such information to improve benefit design where patient cost-sharing for cancer drugs would not simply be based on cost, but value. Value-based insurance design, where co-pays are reduced for services when the clinical benefits exceed the cost and raised when the benefits do not justify the expenditure, has already been implemented by pioneering private payers to drive value in areas outside of oncology (de Souza et al. 2012). As applied to oncology, this more nuanced approach would protect individual freedom of treatment choice, while maintaining controls over the total financial expenditure of health plans, and acknowledge and respond to patient heterogeneity across the entire spectrum of cancer care (de Souza et al. 2012; Drummond and Towse 2012; Pearson and Bach, 2010). Research by Wong et al. (2010) suggests that patients may be willing to pay higher co-pays for more effective treatments.

Fourth, simply making CER evidence available on the relative effectiveness (and costs) of cancer treatments will not be sufficient to impact clinical decision making, especially in the absence of national policy or guidelines requiring implementation of such evidence in practice. The development of decision-support systems and aids, using available evidence of value, could be used to help oncologists make decisions; some available programs can calculate the cost-effectiveness of a given intervention and help physicians better elicit patient treatment preferences (Garber and Sox 2010). Such tools could also be used to help oncologists to communicate with patients about the risks and benefits (and costs) of

available treatment options, including end-of-life care, and realistic expectations regarding their prognosis and response to therapy. Available evidence suggests that oncologists infrequently speak early and openly to patients about their options, the possibility of death, and ways they can make the transition to the prospect of dying (Huskamp et al. 2009; Keating et al. 2010; Wright 2008).

Greater patient-physician discussion of transitions to end-of-life care has been shown to correlate with significantly lower intensive care unit admissions, use of aggressive care, and costs, while promoting patient acceptance of terminal illness, caregiver quality of life, and reduced feelings of regret, depression, and anxiety (Smith and Hillner 2011; Wright et al. 2008; Zhang et al. 2009). Moreover, existing evidence indicates that patients and their families want and need more information on what is available to those at the end of life and, more generally, how the US health system works in this context, and would prefer that such information come from physicians, not elected officials or politicians (DoBias 2011; Harrington and Smith 2008). Support from insurers and national medical associations could be beneficial for fostering shared physician-patient communication and decision making. For example, ASCO could develop guidance and educational opportunities to assist oncologists in considering and discussing costs with patients, as well as less expensive treatment options that might be more appropriate at the end of life. As a starting point, ASCO now recommends that oncologists educate patients on the financial realities of cancer care (ASCO 2009).

Other steps that could be taken to enhance the value of advanced cancer care include more open and substantive debate about how best to care for terminally ill cancer patients. A recent survey conducted by the Regence Foundation and *National Journal* found that roughly 78% of Americans indicated that end of life treatment and palliative care should be part of the public discourse, and 93% believed that such decisions should assume a top priority for the US health system (DoBias 2011). Moreover, the current use of community-based palliative care appears to fall short of what many patients want (Barnato et al. 2007; Goodman et al. 2011; Goodman et al. 2011). Key issues of discussion would include: Is aggressive drug therapy that marginally extends life at a high price tag the best course of action? If so, what constitutes a clinically important difference in survival – 1 month, 2 months, 3

months, longer? Authorities in France, for example, consider 1 month sufficient, while those in the UK, require evidence of at least a 3 month survival benefit. Would greater value for money be obtained by helping patients more effectively transition to the end of life through high-quality (and earlier) palliative care? Existing evidence suggests that earlier access to palliative care can extend life, reduce suffering and save money (Temel et al. 2010).

In a poignant commentary in *The New Yorker*, Atul Gawande maintains that modern medicine has done well at staving off death with expensive and aggressive interventions, but has been considerably less successful in knowing when to stop and instead focus on improving the days terminal patients have left (Gwande 2010). In addition to greater public discussion on these issues, progress on the latter might be attained through improving coordination of care between usual care and hospice, where both could be more effectively combined into integrated care models; payment reform for cognitive care, such as outlining and discussing advanced directives and other end-of-life planning and counselling; and, additional research on determining the optimal time to transition from curative to palliative care and ways to predict death more accurately.

Finally, aspects of value other than clinical and economic outcomes should also enter into public discourse about advanced cancer care - hope, choice, opportunity of treatment, compassion, convenience, equity, and quality of life can all be considered domains of value for patients (and likely other stakeholders, such as their caregivers, physicians, and society). For example, the importance of considering quality of life was highlighted in recent national poll data, which found that most (70%) of Americans agree that it is more important to enhance quality of life for seriously ill patients, even though it means a shorter life (DoBias 2011). Of course, how value is defined, particularly from the patient point of view, will depend on personal circumstances and preferences and what therapies are available. For some patients, this will entail trying every available treatment to prolong life; for others, there may be greater value placed on simplifying their choices and eliminating unrealistic treatment options, while allowing them to live out the remaining months or years of their lives without aggressive therapy, and the health, emotional, and financial implications that may result. Therefore, an important point of discussion is how to

build consensus about when to give and when to omit particular treatments, while still allowing for personalisation of cancer care.

While the aforementioned steps may go some way to ensure the US is achieving greater value in advanced cancer care, there remain broader challenges around the acceptability (and sustainability) of the high prices of cancer drugs and associated services and how much society is willing to spend on treating a small minority of patients and what we might forgo as a result. Resources are limited, especially given the current economic climate, and difficult decisions are required regarding the availability and affordability of health care. While this is an enduring problem, it is preferable to tackle it head on using the best available evidence and determined in a transparent, representative, and fair way.

Study 8: The politics of comparative effectiveness research in the United States: Lessons from recent history

Introduction

As part of the ACA of 2010, the PCORI was established to fund and oversee CER at the national level. As a concept, CER has garnered support from a variety of stakeholders, including government officials, researchers, professional associations, and patients. With growing concerns over health care costs and the quality of care, determining which therapies, care strategies, delivery models, and public health programs are most effective makes good sense. Existing estimates suggest that less than half of medical care in the US is based on or supported by evidence about its effectiveness, often resulting in care that is inappropriate and unnecessary (IOM 2007; McGlynn and Brook 2001).

However, in the year preceding passage of the ACA, the initiative sparked substantial controversy. Many Republicans, private institutions, and conservative pundits went into “rhetorical overdrive” (Iglehart 2010), with claims of government interference in patient care and rationing of services. Town hall meetings resonated with concerns over the creation of “death panels” and fear that the US would adopt a “British-style” model of health care. Despite intense opposition and attempts to gut the funding in the final reform package, CER managed to survive the legislative process and is now well under way. Nevertheless, these first years of formulating and implementing CER represent a challenging new chapter in the initiative’s history. Many unresolved questions remain, including how CER will be implemented and operationalised in practice and, ultimately, what impact it will have on its intended aims to improve the quality of health services, eliminate inappropriate and wasteful care, and “bend the cost curve.” There is also uncertainty regarding its survival in both the near and the long term. With ongoing attempts to overturn provisions in the health care reform law, the debate regarding CER is far from over.

Although CER has garnered significant attention in recent years, the issues and methods underpinning the approach are not entirely new. There have been a number

of efforts at the federal level over the last forty years, some more successful than others, to improve the evidence base on the benefits and costs of new health care interventions and policies. Therefore, the aim of this article is to examine past attempts to implement and use this type of research, in order to identify lessons for current and future CER efforts. To date, there has been a lack of research on previous attempts at CER and the implications these efforts have had or might have for current CER initiatives. The article begins by providing further policy and political context around recent CER legislation and the creation of PCORI, followed by a historical analysis of previous efforts to formally integrate related research into US health care. It then turns to a discussion of key lessons from past attempts and concludes by highlighting two main challenges CER will likely face in the future. We suggest potential ways to address these pitfalls to ensure the success and sustainability of CER.

Setting the policy context

Although debate around CER came to a head during formation of the health care reform bills during the summer of 2009, discussions on its potential role in health care largely took shape earlier, in late 2006 and early 2007. During this time, a shift from a Republican to Democratic congressional majority renewed focus on health care and the need for comprehensive reform (Patel 2010). In this context, CER was increasingly viewed as a mechanism to address deficits in the quality and efficiency of health care by better integrating available evidence on effectiveness into care delivery (CBO 2007). There were also growing calls among congressional leaders and health policy experts for an identifiable entity responsible for the research (CRS 2007; IOM 2007; Schoen et al. 2007; Wilensky 2006).

Some of the key policy questions fundamental to these discussions were issues related to the governance, structure, and funding of such an entity. Of particular importance were whether it would be connected with or independent of government and what role the research would have in Medicare coverage decisions and other areas of policy, such as the development of clinical practice guidelines. A proposal put forth by the House in 2007, for example, introduced a potential organisational

structure that would be housed within the AHRQ. Alternatively, Senators Max Baucus (D-MT) and Kent Conrad (D-ND), chairs of the Finance and Budget Committees, respectively, introduced legislation (Comparative Effectiveness Research Act of 2008, S. 3408, 110th Cong. 2008) that would establish a nongovernmental, public-private entity, based on the idea that a nongovernmental body would offer a more efficient, transparent, and accepted mechanism for integrating CER into US health care. Although neither proposal was fully adopted, these broader questions about the structure and role of a CER entity and the application of such research in policy and practice remained central to subsequent health care reform debates and the final reform bill.

During the 2008 presidential election, momentum for and interest in CER continued to grow. Both major party nominees made CER a key component of their approach to health care reform. Although CER garnered bipartisan support in the campaign, once Obama was elected in November 2008 and the Democrats took the election, Republican support waned, with many party leaders distancing themselves from or opposing the issue (Iglehart 2010). As the battle over health care reform intensified and the stimulus bill was being debated in the early months of 2009, the discussion shifted from the specifics of CER and a possible research institute to whether it should be included in the reform at all (Patel 2010).

During this time, opponents of CER found their voice in conservative commentators that linked the research with government domination over health care and warned of the negative implications of adopting foreign models of CER, such as the NICE in the UK (Avorn 2009; Conservatives for Patients' Rights 2009; Iglehart 2010). Such opposition centred on fears that CER would restrict patients' access to care and physicians' autonomy and threaten biomedical innovation (Iglehart 2010). For example, Senator Mike Enzi (R-WY), the ranking Republican member of the Senate HELP Committee, asserted that a CER institute would "create a new bureaucracy to dictate which treatments to pay for" (US Senate Committee on HELP 2009). Parallel arguments were made by political pundits and, perhaps not surprisingly, industry representatives (Will 2009; Zhang 2009).

As health care reform legislation worked its way through House and Senate committees during the latter part of 2009, Republicans called for dozens of amendments focused on eliminating funding for a research entity, restricting considerations of cost or cost effectiveness in the research, and prohibiting research results from being used in coverage decisions. Both committees received an endless stream of comments from organisations, researchers, and citizens expressing concerns that CER would dictate insurance coverage decisions, prevent Americans from receiving innovative therapies, and hamper developments in personalised medicine. Committee staffs, in turn, gathered statements of support from various stakeholders and launched an educational campaign of sorts to clarify intensifying misperceptions that CER would deny or dictate care (Patel 2010). Despite these distractions, congressional leaders successfully guided their respective versions of comparative effectiveness entities through the legislative process: the House called for a government based institute, while the Senate supported a public-private entity.

In the process of reconciling the two bills into a final reform bill, the concept of a public-private institute—PCORI—prevailed. In compromise with the House, funds would be dedicated to AHRQ for dissemination of the research. Moreover, a number of calculated decisions and negotiated amendments were made to garner bipartisan and stakeholder support. For instance, the bill language prohibited the use of QALYs, a metric used in cost-effectiveness to measure net health gain, as well as the use of research findings to dictate coverage, reimbursement, or other policy recommendations (Neumann and Weinstein 2010). This reflected the need not only to counter any claims that health care reform would lead to rationing, but also to protect PCORI from perceptions that it was sponsoring research to be used by Medicare for coverage and reimbursement decisions. The bill also emphasised commitments to transparency and stakeholder representation in almost every aspect of the institute's operations, which were central tenets of the House bill. Such actions were intended to allay fears of government takeover of health care and maintain independence of the scientific process (Iglehart 2010).

In the end, PCORI was charged with identifying research priorities, establishing a research agenda, and coordinating CER. To realise these aims, the

institute draws on a dedicated trust fund of dollars from the Medicare program and contributions from private insurers, with initial expectations of an annual budget of up to \$500 million within a few years (Iglehart 2010). As previously discussed, substantial funding is dedicated to AHRQ to disseminate CER findings and link databases and disease registries to improve evidence generation.

Past Federal attempts to formalise CER and related approaches

To better understand the potential challenges and opportunities facing the CER initiative, we now explore key past attempts to establish a formal role for CER and related research approaches (Table 13). Assessing the history of related efforts may offer important lessons, considering that US health policy, especially on contentious issues, often develops incrementally: it is often tried, modified or scaled back, or sometimes dropped before it is eventually adopted. Indeed, current CER policy can be viewed as the latest in a long line of related proposals, policies, and initiatives that started in the 1970s.

First, however, it is useful to outline the different terminology commonly used to characterise CER and related activities. Although terminology has changed over the years, the underlying concepts remain quite similar between HTA, evidence-based medicine (EBM), and CER. In the 1970s, such efforts were mainly framed under the rubric of HTA. However, throughout the 1980s and 1990s, the term HTA was often replaced by terms such as outcomes research and effectiveness research, and later, in the 2000s, with EBM and CER. Luce et al. (2010) note that much confusion remains around how these terms—namely, HTA, EBM, and CER—are used or intended and offer a framework to lend greater clarity. Both HTA and CER address the question “Does it work?” and involve evidence synthesis. However, they can be distinguished by the fact that CER also focuses on evidence generation and is principally concerned with the comparative assessment of effectiveness of a broad range of interventions and care delivery approaches in routine settings, whereas HTA considers evidence on effectiveness, safety, cost-effectiveness, and, when broadly applied, social, ethical, and legal aspects of health technologies. Because HTA often (but not always) includes an economic dimension (cost-effectiveness), it also addresses the question “Is it worth it and should it be paid for?” and is often used to

inform coverage and reimbursement decisions. As previously discussed, these are aspects that are not included in current conceptions of CER. Similar to HTA and CER, EBM is underpinned by evidence synthesis (largely via systematic reviews), but differs in terms of also entailing a decision process, where such evidence is used to support individual patients' and physicians' clinical treatment decisions. Taken together, CER can be viewed as a potential input into HTA and EBM.

In the following sections, we primarily reference the terms that were used at the time a particular initiative was conceived or adopted. While there is no widely accepted definition of these terms, the framework outlined by Luce et al. (2010) provides a guide for understanding the underlying focus of different approaches to ascertain the benefits and costs of health care interventions and policies.

Office of Technology Assessment (OTA)

One of the first key federal efforts was the OTA, which was established by Congress in 1972 at the request of Senator Edward Kennedy (D-MA) to serve as an advisory body to Congress on technical and scientific policy issues, including the use of medical technologies. The office was overseen by the Technology Assessment Board, a bipartisan committee of six senators and six representatives drawn equally from the two parties, and only the board or the ranking member of the minority party could request studies. Given its placement in Congress, OTA reports rarely outlined specific recommendations; rather, they summarised the available evidence, identified policy alternatives, and discussed their advantages and disadvantages (Bimber 1996). Consequently, legislators on opposite sides of contentious policy issues often cited the same report to advance their respective lines of argument. As a result of its neutrality and robust analyses, the OTA was highly praised in government, academic, and scientific circles and served as a model for other countries interested in establishing a similar body.

However, after the general election of 1994, both the House and Senate had new Republican majorities determined to enforce fiscal discipline, shrink the size of government, and reduce governmental regulation. Congress decided to start by cutting its own budget - eliminating one of its support agencies, the OTA; cutting the

budget of others; and reducing staff in general (Bimber 1996). Although legislators who supported the agency, such as Representative Amo Houghton (R-NY) and Senators Ted Stevens (R-AK) and Charles Taylor (R-NC), made several attempts to save the agency, it was not enough to shift the vote to withdraw its funding, and the OTA closed in late 1995.

Reasons for the agency's defunding include criticisms over its explicit inclusion of costs and cost-effectiveness in assessments, the fact that its research duplicated the work of other public and private organisations and often lagged behind legislation, and some structural elements that increased its vulnerability (Kunkle 1995). For example, Senator Connie Mack (R-FL), who helped lead the effort to kill the agency, and other opponents maintained that its role could be filled by other congressional agencies, such as the GAO and the Congressional Research Service (CRS), or private entities (e.g., the National Academy of Sciences). Moreover, because the OTA's structure insulated it from direct contact with most members of Congress in attempts to maintain political neutrality, many policy makers had very limited personal knowledge, experience, or appreciation of the agency. Therefore, when the OTA was under immediate threat, there was not a critical mass of support to save it. Moreover, knowledgeable observers believed that some of the OTA's assertions, particularly those pointing to the lack of available evidence to allow rational and objective utilisation of medical technologies, likely alienated organised medicine and the drug and device industries and led them to favour the agency's elimination (Eisenberg and Zarin 2002). Certainly, the device industry's position that "no single provision of health reform could work greater harm on medical innovation or patients than national assessments of technologies before they could be used by local plans" (Tunis and Gelband 1994: 354) supports this observation—and these sentiments were formally expressed again in the 1994 testimony on health care reform. Critics of the closure viewed it as an example of politics overriding science and evidence, and since its closure, proponents of the agency, such as Hillary Clinton, have called for its reinstatement (Healy and Dean 2007; Malakoff 2001).

National Center for Health Care Technology (NCHCT)

In 1978, six years after the OTA was established, the National Center for Health Care Technology (NCHCT), an agency in the Executive Branch and housed within the DHHS, was created to promote and support HTA, to conduct comprehensive assessments of technologies with important national implications, and to advise the Health Care Financing Administration (HCFA) on Medicare coverage issues. The tenure of the NCHCT was brief, and its operations limited, with a small staff (around twenty) and a minimal budget; four years after its inception, the Reagan administration eliminated the agency by zeroing its annual budget and transferring many of its functions to the National Center for Health Services Research (NCHSR), later the Agency for Health Care Policy and Research. Cotter (2009) suggests that the demise of NCHCT was attributable largely to opposition from industry and medical associations, particularly the American Medical Association (AMA), who argued that the “relevant clinical policy analysis and judgements are being made—and are being responsibly made—within the medical profession” (Boyle 1981: 297) and therefore the agency was redundant and overstepping the role of clinicians. The AMA was also in strong “opposition to those who would make cost-effectiveness the essential factor in determining whether medical care is . . . reasonable and necessary” (Boyle 1981: 297). In addition, industry representatives often attributed the center’s work as attempts to regulate industry and bluntly cut costs—a likely reaction to the fact that about 40 percent of its recommendations were for non-coverage of evaluated technologies. The small operations of the center were also nested in the DHHS, whereas more powerful, affluent, and prominent agencies, such as the NIH, had much greater departmental attention and support and much stronger constituencies (Blumenthal 1983).

Council of Health Care Technology (CHCT)

After eliminating the NCHCT, Congress still perceived a need for some capacity to evaluate health technology. In response to a congressional mandate, the Council on Health Care Technology (CHCT) was established in 1986 within the Institute of Medicine (IOM) to promote the concept and use of HTA and review health technologies for their appropriate use. It was considered the first public-

private technology assessment partnership in the US (Luce and Cohen 2009), receiving “matching” government financial support on the condition that it first secure private funds. In the end, the CHCT focused primarily on conceptual and methodological issues in technology assessment (e.g. approaches to priority setting, and the relationship between technology assessment and quality assurance), producing only two technology assessments during its operation (OTA 1995). The council’s lack of tangible output brought its usefulness into question. Moreover, from the beginning its goals were never clear, and the need to raise private funds hampered its operations. The IOM therefore did not seek further public funds for the council in 1989, and its statutory authorisation was allowed to expire.

Agency for Health Care Policy and Research (AHCPR)

Soon after the council’s demise in the late 1980s, policy makers became increasingly concerned with health care costs and the financial sustainability of Medicare. There was also worry that the new prospective payment system for Medicare inpatient care was forcing patients out of the hospital earlier than clinically warranted because their “DRG had run out” (AHRQ 1999). At congressional hearings on Medicare, the phrase “quicker but sicker” reflected a central concern about the implications of such new financial incentives for quality of care. Concurrently, the DHHS convened several meetings, attended by John Wennberg, William Roper, David Eddy, Robert Brook, and others, to explore whether outcomes or effectiveness research could be useful on a large scale to define optimal treatments and monitor and improve quality of care. In these meetings, Wennberg (1984) emphasised that the necessary scientific information was needed to allow physicians to define optimum treatments. Influenced by Wennberg and colleagues, several Republican and Democratic congressional leaders heralded the idea that evidence-based approaches, namely outcomes research and technology assessment, could potentially save billions of dollars by identifying unnecessary health care services in high-cost areas without harming patients (Gray, Gusmano and Collins 2003). The AMA also expressed interest in and commitment to reducing “waste” through scientific study and evidence-based guidelines.

Subsequently, the Agency for Health Care Policy and Research (AHCPR) was created in 1989 to carry out outcomes studies, develop practice guidelines, and conduct and coordinate health services research. While its establishment attracted no opposition, there was a lack of widespread agreement among policy makers and stakeholders that federal support for effectiveness and outcomes research was entirely warranted.

In the first few years of its operation, the new agency was by several measures successful. Under its Medical Treatment Effectiveness Program, it established a new program for developing practice guidelines as well as fourteen new Patient Outcomes Research Teams (PORTs), multidisciplinary centers focused on particular medical problems (e.g. back pain, myocardial infarction) to determine “what worked” (Gray, Gusmano, and Collins 2003; Luce and Cohen 2009). The agency also established the Office of Health Technology Assessment (OHTA), which conducted formal, comprehensive assessments, often at the specific request of Medicare (Luce and Cohen 2009). Furthermore, the AHCPR collected and analysed large amounts of health services data through the National Medical Expenditure Survey, which was used to assist the White House, under both the Bush and Clinton administrations, in assessing the implications of different health care reform policy options.

The agency’s successes in implementing its legislative mandate led to substantial budget increases, from \$53 million in 1989 to over \$162 million by 1995. However, in 1994, following introduction of the conservatives’ broad government reduction agenda (notably the Contract with America), the AHCPR’s performance underwent significant scrutiny, and the agency was almost eliminated during the 1996 budget appropriations. Its vulnerability to termination was attributable to many factors, several of which relate to the strategies underlying its creation (Gray, Gusmano, and Collins 2003). Namely, evaluations of the agency focused on its inability to meet the original expectations of Congress, deeming its practice guideline program ineffective and expressing doubts that its effectiveness research work (via PORTs and use of administrative databases) would result in cost savings (GAO 1995; OTA 1994; PPRC 1995). Moreover, as with other preceding agencies, some of its work was viewed as redundant with that of other federal agencies and private

organisations, and its occasional use of cost-effectiveness information in its technology assessments and practice guidelines often generated controversy.

The opposition of influential lobby groups also had a substantial impact. In particular, the North American Spine Society (NASS), an association of back surgeons, mounted an attack on the agency in 1995, with the support of a number of Republican politicians, after its PORT on low-back pain concluded that “there was no evidence to support spinal-fusion surgery and that such surgery commonly had complications” (Gray, Gusmano, and Collins 2003: W3-297). The AHCPR ultimately survived the backlash, thanks to a series of behind-the-scenes negotiations between the NASS and the former members of Congress who helped create the agency. These supporters emphasised that many of the problems that underpinned its establishment (e.g. practice variations and cost concerns) remained and therefore the agency was needed. Nevertheless, there were repercussions. The agency gained a new name, the Agency for Healthcare Research and Quality (AHRQ), a deliberate move to remove the word *policy* from the name; its practice guideline program was abandoned; and it received a 21 percent budget cut (Gray, Gusmano, and Collins 2003).

Support for the agency once again gained momentum in the early 2000s: in 2003 the Medicare Modernization Act (MMA) authorised the AHRQ to strengthen its role in conducting and disseminating CER to “improve the quality, effectiveness, and efficiency of health care” delivered to Medicare, Medicaid, and State Children’s Health Insurance Program (SCHIP) enrollees. The Effective Health Care (EHC) program was established to meet this aim, which oversees a range of external research networks that conduct systematic evidence reviews to assess the effectiveness, comparative effectiveness, safety, and, in rare cases, cost-effectiveness of medical technologies and interventions (Sullivan et al. 2009). While the MMA demonstrated the continued interest among policy makers in comparative information on the value of health care interventions, it also highlighted the strong opposition to provisions that would “strengthen the hand of government” over industry (Neumann et al. 2005). In particular, the legislation explicitly prohibited the AHRQ from mandating national standards of clinical practice and banned the Centers for Medicare and Medicaid Services (CMS) from using comparative

effectiveness information produced by the EHC to withhold or restrict access to prescription drugs (Neumann et al. 2005; Sullivan et al. 2009). Moreover, the agency's budget was once again reduced—the originally authorised \$50 million per year for CER activities was ultimately reduced to \$15 million.

Medicaid and Medicare

Other federal efforts to formalise the role of CER-related activities were introduced in the mid-1990s and early 2000s, particularly with regard to Medicare and Medicaid. In 1994 Medicaid, and in particular the Oregon Medicaid program, sought to create a list of medical services, ranked according to cost-effectiveness analyses and public preferences (elicited through surveys and community meetings), to help determine which services would be covered under the Oregon Health Plan (Oberlander 2007; Oberlander et al. 2001). The core idea was to use any savings from rationing health services (by not covering low-ranking services) to expand Medicaid coverage across the state. The initial list and approach to priority setting received significant national controversy on political, technical, legal, and ethical grounds (Brown 1991; Eddy 1991; Oberlander et al. 2001). Consequently, state administrators were required to make political concessions (e.g., cost-effectiveness analysis no longer used) and to modify the list (Blumstein 1997). Oregon still uses some elements of this approach today, but perhaps unsurprisingly, it has not been adopted by any other state Medicaid program. However, several states do participate in the Oregon Drug Effectiveness Review Project (DERP), which provides state Medicaid agencies with comparative effectiveness and safety reports to inform their drug formulary decisions. Still, there is no shortage of controversy surrounding the DERP program (Neumann 2006). Pharmaceutical manufacturers have criticised the program's reports as providing "political cover" for cost-containment decisions taken by state Medicaid programs. Others have expressed concern that DERP's selection of evidence is too strict, thereby restricting the use of all possible available research that could inform coverage policies (Steinberg and Luce 2005).

Like Medicaid, Medicare has a complicated history with CER. Stipulated in authorising legislation in 1965, Medicare pays for medically necessary services provided to elderly and disabled individuals. For many years, "reasonable and

necessary” was understood to reflect the prevailing views of physicians, although there were no formal criteria to define this standard at either the local or the national level (Neumann et al. 2005). Amid growing complaints that its coverage process was unpredictable, unclear, and lengthy, in 1989 the HCFA (later the CMS), under the Reagan administration, pushed regulation stating that for purposes of coverage, a technology would need to be safe, effective, non-investigational, appropriate, and accepted by the medical community (HCFA 1989). The inclusion of costs and cost-effectiveness as an explicit criterion in selected cases was also proposed, and this was the first time that HCFA had supported such considerations as criteria in coverage decisions. The proposal generated substantial opposition from medical and industry groups (e.g. the AMA and the Pharmaceutical Research and Manufacturers Association of America), who feared that important technologies would be rationed, leaving seniors to pay for or forgo necessary care, and was eventually withdrawn (CMS 1999; Foote 2002; Pear 1991).

Almost a decade later in 2000, during the Clinton administration, the CMS issued a notice of intent (NOI) to publish a proposed rule in which it did not mention cost-effectiveness per se, but did mention the concept of “added value” (CMS 2000). The proposed standards would require that new technologies provide some benefits to beneficiaries beyond what was already available to them in the program. If a technology did not provide added value, Medicare would deny coverage. The costs of alternative treatment options were, once again, considered relevant to determining coverage. However, a significant number of negative comments on the NOI persuaded the DHHS and the CMS to withdraw this proposal, and no further attempt has been made to address these issues through regulation.

In a number of ways, Congress has made the CMS’s interest in considering costs and comparative value more difficult, suggesting that these occasional efforts to develop national policies that explicitly address costs or cost-effectiveness have, in fact, resulted in policy moving in the opposite direction (Keenan, Neumann, and Phillips 2006). For example, when the CMS attempted to pay a single rate for products deemed “clinically equivalent,” Congress expressly prohibited the agency from any future use of this standard involving payments to hospital outpatient departments (Neumann, Rosen, and Weinstein 2005). Moreover, both the

Congressional Budget Office and the Medicare Payment Advisory Commission have asserted that regardless of legal interpretation of the current statute, the CMS would require clear statutory authority to formally consider costs in its coverage policies (CBO 2007; Medpac 2008). That authority is unlikely to be given, at least in the short term. As discussed earlier, the ACA prohibited Medicare from using a specific threshold for a cost per QALY in making reimbursement and coverage decisions, and limitations were placed on the ways CER could be used.

Another issue hindering the CMS from more frequently considering comparative effectiveness evidence in coverage decisions is that there are significant gaps in the available evidence bases, particularly as related to seniors and disabled Medicare beneficiaries. In part to address this issue, in 2006 the CMS issued guidance on a new policy called coverage with evidence development, which links coverage of an intervention with a requirement that patients receiving it be enrolled in prospective clinical studies to inform future revisions to the coverage decision (Tunis and Pearson 2006). The aim is to allow beneficiaries access to potentially promising technologies where available evidence on their clinical effectiveness is either insufficient or restricted to a particular patient population. Unlike previous attempts to codify the role of evidence in the CMS's coverage policies, CED has largely escaped political backlash through its framing as a mechanism to *enhance* the adoption of new therapies, rather than to control or slow their diffusion (Luce and Cohen 2009), and because it avoided the issue of cost-effectiveness. Moreover, the policy has been applied in a limited number of cases since its inception, and only in two cases have the resulting evidence been used for policy, thereby lessening any perceived threat. Yet the policy has faced some opposition from patient advocacy groups as well as from manufacturers, who argue that CED raises the threshold of evidence needed to obtain a positive coverage determination and slows access to medical advances (Robinson 2010). There is limited evidence to support such concerns, however. For example, despite available comparative effectiveness evidence demonstrating that computed tomography of the coronary arteries offers little advantage over current angiographic approaches, the CMS has not been able to stop paying for the procedure due to strong resistance by radiologists (Appleby 2008; Redberg and Walsh 2008).

Veterans Health Administration (VHA)

The VHA has been conducting CER and putting the findings into practice for many years, beginning in the mid-1990s (Kupersmith 2009). There is no single VHA office that conducts HTA; rather, several in-house groups assume this role, conducting various forms of HTA with different purposes across the administration (Luce and Cohen 2009). Unlike other past CER efforts, the VHA has been successful in going beyond effectiveness research only to include costs and cost-effectiveness, and evidence is directly linked to coverage, acquisition, and care delivery decisions (Luce and Cohen 2009). A centralised staff-model structure enables researchers to access data from a variety of sources, and researchers have access to a clinical research network. There are also mechanisms in place for information dissemination, allowing the results of comparative effectiveness analyses to be broadly released within the VHA system. The centralised structure also facilitates monitoring of compliance with any new coverage or payment rules based on CER results.

Other elements of the system have contributed to the success of its CER activities. For instance, because the VHA is essentially an integrated, closed system that operates within an appropriated budget constraint, it faces pressures to offer the most effective range of health services from available resources. It also has a significant purchasing arm that exercises substantial control over providers and patient care. Furthermore, its CER activities are embedded in a history of reforms and strong leadership that have placed considerable emphasis on quality improvement (Oliver 2007). This has led to investments in performance management, a sophisticated health electronic record system, and VHA-funded health services research—policies that demand for and utilise research to evaluate health outcomes and costs and, subsequently, identify areas where quality of care can be enhanced.

In sum, recent CER initiatives bear resemblance to previous attempts to support similar evidence-based approaches at the Federal level. However, most of these efforts failed to get off the ground, were defunded or faced defunding, and/or were downscaled significantly. While the actual performance of these efforts, in

terms of quality, objectivity, transparency, and relevance, came into question on occasion, other factors, mostly political in nature, also significantly contributed to their lack of success, including perceptions that the research would impose barriers to technology innovation and delay market access, provide a guise for government control on the practice of medicine, and place cost considerations above patient care.

Table 13: *Summary of past Federal attempts to formalise CER and related approaches*

Agency/Organisation	Tenure	Funding	Governance Structure	Remit	Approach	Major Accomplishments/Failures
Office of Technology Assessment (OTA)	1972-1995	\$20M	Overseen by the Technology Assessment Board, a bipartisan committee of Senators and Representatives (equal representation).	Advise Congress on scientific and policy issues related to technology; conduct studies and publish reports on technical and scientific issues.	Assessments included considerations of safety, effectiveness, costs, cost-effectiveness, and wider social-economic benefits and costs.	Produced objective and robust reports that were used to support policy across party lines. Issued over 750 reports. Agency and its work were not visible enough to a broad spectrum of policy makers and other stakeholder groups. Also, did not distinguish its role and contribution from that of other governmental agencies, and reports were not always timely enough to influence policy. May have alienated organised medicine and drug/device industries.
National Center for Health Care Technology (NCHCT)	1978-1982	\$4M	Part of the Executive Branch, housed within DHHS.	Assess high-impact technologies and advise on Medicare coverage.	Assessed safety, efficacy, effectiveness, costs, and cost effectiveness, social, ethical and economic impacts.	Established extramural research grants program on technology assessment and assessment methods, performed technology assessments itself, and evaluated about 75 technologies for coverage by Medicare. Also, frequently involved medical professionals and industry groups collaboratively in its work. Agency was not able to garner industry and clinical community support; small, with limited staff; disappeared within the DHHS; no appropriations given in 1982.

Council on Health Care Technology (CHCT)	1986-1989	\$750K (max, 3 rd year of operation); contingent upon matching support from non-Federal sources.	Part of the Institute of Medicine; involved health care and health economics experts, health professionals, insurers, patients).	Promote HTA and HTA methods and assess health technologies.	Unclear (conducted few studies).	Enhanced awareness of HTA and created various resources on assessment activities in US, HTA methods, and individual assessments. Unclear mission/goals and lack of tangible output. Also, inability to raise sufficient private funds to continue operation.
Agency for Health Care Policy and Research (AHCPR)	1989-present	\$52M (1989) \$390M (2011)	Agency of the Department of Health and Human Services.	Carry out outcomes research, develop practice guidelines, and conduct and coordinate health services research.	Assessments include considerations of safety, effectiveness, costs, cost-effectiveness, and wider social-economic benefits and costs.	Productive clinical guidelines program; supported use of health services and outcomes research in devising health reforms; secured strong bi-partisan support base to guard against opposition in critical moments. Inability to demonstrate impact on health care practice and cost-savings; problems showcasing value of work compared to similar Federal and private efforts; occasional reductions in funding and removal of certain programs.
Medicaid (Oregon)	1994-present		Agency of the Department of Health and Human Services.	Oregon Medicaid: Create a list of covered medical services based on evidence. DERP: Provide state Medicaid agencies with CER and safety information.	Uses evidence from cost-effectiveness analyses, surveys, and community meetings to rank the value of available medical services.	Pushed agenda on the need to consider cost-effectiveness in policy and to eliminate coverage for low-value services. Conceded to political pressures and dropped incorporation of cost-effectiveness evidence; lack support from certain stakeholder groups; approach limited to Oregon.

Medicare	Late 1980s-present*	No dedicated funding for CER.	Agency of the Department of Health and Human Services.	Provide access to health insurance for the elderly, younger persons with disabilities, and those with end stage renal disease.	Considers whether a service is “reasonable and necessary”, which may take into consideration comparative effectiveness; costs and cost-effectiveness is not considered.	Established the CED program. Inability to include costs and cost-effectiveness, despite efforts to do so; legislative restrictions on use of evidence in Medicare policy; limited applications to CED to date.
Veterans Health Administration (VHA)	1994-present	No dedicated funding for CER.	U.S. Department of Veterans Affairs.	Provide comprehensive health care to US Veterans.	Assessments include considerations of safety, effectiveness, costs, cost-effectiveness, and wider social-economic benefits and costs.	Formally considers a wide range of factors, including costs and cost-effectiveness, in its research and uses evidence in decision making; robust infrastructure for CER and dissemination of evidence; well-established organisational culture supportive of CER and quality of care improvement.

Source: Authors' compilation

*When the first proposals to consider CER evidence in Medicare policy were introduced.

Given the substantial investments made to date in CER and the uncertainty surrounding its ultimate role in and impact on health care, it seems prudent to elucidate what previous initiatives have taught us, which we discuss in the following section.

Lessons learned

First, past experiences suggest that while the concept of CER may be widely accepted in principle, support may wane if resulting evidence challenges current policy and practice. For example, in both the US and abroad, a public backlash has occurred when available evidence calls for some degree of disinvestment or places conditions on access to or payment for care. In such cases, the research (and general overall approach) is seen as a potential threat to the interests of vested stakeholders, namely organised medicine, industry, and patient advocacy organisations—groups who typically possess significant resources to influence policy makers (Gerber and Patashnik 2010).

This was typified in the cases of the NCHCT and AHCPR. In the former, technology manufacturers, supported by organised medicine, expressed extreme consternation over the potential threat that the center's list of "emerging" medical innovations for evaluation would have on investments in innovative medical technologies (Luce and Cohen 2009). In turn, these groups raised doubt surrounding the NCHCT's authority and utility by pointing out that it was unduly regulating industry and duplicating assessments carried out by physicians, leaving it vulnerable to termination during the Reagan administration's government reductions. Opposition from organised medicine, namely the NASS, was similarly instrumental in the AHCPR's near-death experience. The AHCPR case also illustrates that the prevailing political context can have an influential impact on the success of interest group lobbying efforts. Other AHCPR studies had drawn criticism by medical associations in previous years, but did not result in attacks on the agency by members of the administration or Congress (Gray, Gusmano, and Collins 2003). The efforts by the NASS were more successful largely because the House Republicans, who enjoyed a majority in the House, were prepared to believe the worst about the

agency, given that such criticisms aligned with their ideological assumptions about the government's performance.

Second, many people who supported the creation of previous agencies to carry out CER-related activities often heralded unrealistic aims about what the research could achieve and within what time frames. For example, the members of Congress who championed the creation of the AHCPR believed that outcomes research, HTA, and practice guidelines would reduce health care spending and alleviate Medicare's financial problems, and they justified supporting federal funding on this basis. Congressional opponents later used the fact that the agency's research did not substantially save money as evidence that it had no impact and was thus a waste of tax dollars. It is possible that the AHCPR would have achieved greater impact, especially on costs, over the longer term, but by the mid-1990s the political tide had turned against government influence in the health sector, emphasising the challenges raised by the brevity of political time frames.

Third, previous initiatives in the US point to the importance of "political champion(s)" or "political entrepreneur(s)" to gain support for CER, especially in terms of its use in policy and practice. Without reinforcement from influential congressional leaders, the AHCPR would likely have folded under attack from conservative opposition and organised interests. The OTA, however, was less fortunate. Although the agency received widespread support for its reports and was almost saved by congressional supporters when it faced defunding, its arm's-length position and lack of visibility among a range of policy makers left it vulnerable to elimination. The OTA case suggests not only that it is important to have political champions, but also that such support must be sustained over time, visible to a variety of stakeholders, and, ideally, bipartisan in nature.

Fourth, previous organisations often suffered from a lack of a single, authoritative "voice" in defining research agendas and offering guidance and recommendations. Consequently, it was unclear how their role differed from other public and private CER efforts, leading to questions about value. To illustrate, one of the reasons agencies such as the OTA, the AHCPR, and the NCHCT were targeted

for elimination was that their work was viewed as duplicative of existing initiatives or programs and thus unnecessary. This was often exacerbated by weak or unclear articulation of their aims and remit.

Fifth, explicit consideration of costs has often placed agencies at threat by opponents, particularly those against the use of cost-effectiveness analysis in policy making. The OTA, the NCHCT, and some Medicaid and Medicare programs or provisions were eliminated, came under threat, or were significantly modified in part because they explicitly considered (or proposed considering) costs and cost-effectiveness in their analyses and recommendations. Doing so effectively increased their susceptibility to negative media coverage, interest group and lawmaker attacks, and public debate around rationing and government overreach when unfavourable decisions or recommendations were made, or when critics needed additional justification to eliminate a particular agency or program (Oberlander and Marmor 2001). In reaction to criticisms, political concessions often followed to disregard cost or cost-effectiveness evidence in a specific context or to drop their consideration in policy overall, which one can argue placed these agencies at further risk for attack by hampering their ability to achieve some of the aims they sought to achieve (e.g. reduced costs).

Finally, in some ways, one of the most significant lessons is that the incentives inherent in the American health system have helped deter the success of past CER efforts. For example, the lack of infrastructure, requirements, and legislative mandates for use of CER in policy making have limited its impact on health care delivery and costs. Providing evidence alone has not been sufficient to meet the aims that CER initiatives have intended to achieve. In addition, because of health financing practices in the US, where volume of care rather than value is rewarded, providers have had limited incentive to consider comparative effectiveness, quality, or costs of care in their practice decisions. Furthermore, private insurers have faced little competitive pressure to attain greater value by improving quality and/or reducing costs; rather, the cost of new interventions is passed along in premium increases and limited benefits (Kaiser Family Foundation 2012). While the growing unaffordability of health insurance has received attention in recent years, insurers

have largely escaped scrutiny over the years due in part to the limited voice of the insured population. Moreover, clinical decisions are rarely subjected to any formal review of quality or adherence to up-to-date standards of care, which further contributes to making the use of evidence in practice discretionary.

Discussion

Given the past efforts to identify what works in health care through CER-related activities, what can or should be done to help support the success and sustainability of current CER initiatives? In many ways, the architects of health care reform have potentially protected PCORI from some of the political and practical challenges faced by previous agencies through its design features. For instance, linking PCORI's financial support to a separate fund derived from both public and private revenue streams will help protect it from the whims and unpredictability of the annual budget process. Moreover, the fact that the institute is independent from government may go a long way in providing an offense and defense against potential claims that its work is simply a guise for government control over health care decisions. The legislation also provided for broad stakeholder involvement in PCORI, from representation on its committees to offering opportunities for public input on its processes, methods, and prioritisation of comparative effectiveness studies. Together with an emphasis on transparency and incorporation of conflict of interest protections, such provisions may serve to insulate the institute and CER more generally from uneven stakeholder influence.

The fact that PCORI's leadership and committees contain a diverse set of representatives from multiple federal agencies and other organisations may also help improve coordination across different CER-related activities, lending to protection against accusations that it is merely duplicating other existing public and private efforts. In addition, prohibiting analysis of costs and cost-effectiveness and direct links between research findings and policy, namely coverage determinations, will likely render the initiative more politically palatable in both the short term and the long term. To that end, although PCORI does not have an "authoritative voice" over direct applications of CER in policy and practice, it does have significant influence

over determining the research agenda, methodological standards, and in funding both public- and private-sector comparative studies.

Finally, CER has benefited from support from a range of champions on both sides of the political divide. Senator John McCain (R-AZ) and other prominent Republicans backed CER during the 2008 campaign, and President Obama has activity presented CER as one of his central policy initiatives (Frederick 2008; Obama-Biden Presidential Campaign 2008). However, it is uncertain whether such support can be sustained when there is a change in leadership. For example, during the 2012 presidential debates, former Massachusetts governor Mitt Romney repeated the charges heard during earlier health care reform debates that the ACA represented a government takeover of health care, and Republican pundits expressed familiar rhetoric historically used to describe CER, focusing debates on rationing, restricted access to care, and threatened innovation.

Based on past history and the evolution of CER to date, we anticipate two major pitfalls that CER may face in the future. First, in principle, most stakeholders support the concept of more and better information about what works and what does not in health care. However, in cases where CER might contradict or change current practice, especially if it entails restrictions on patient access or provider choice of care, such support will likely wane, and entrenched stakeholder interests and agendas may once again rise to the surface. The example of the 2009 US Preventive Services Task Force guidelines on mammograms highlights that even with removing mandated links between comparative effectiveness studies and policy decisions, the application of research findings may still prove contentious. These recommendations, which advised against biennial screening in women aged forty to forty-nine years, were attacked by a broad range of stakeholder groups, and in response the ACA included a provision requiring the federal government to ignore them (Gusmano and Gray 2010). While stakeholders arguably had legitimate concerns about the validity of applying aggregate results to individual patients, most of the attacks on the guidelines ignored the nuances of the recommendations and accused the government of placing cost savings ahead of the lives of women. Therefore, whether CER will ultimately play a meaningful role in health care decision making and delivery will depend on the willingness of patients, providers,

industry, and other stakeholders to accept limits on access to certain technology and interventions, particularly those with limited or questionable value.

Given these challenges and the experiences of past initiatives, it will be important to continuously promote the value of CER over time. It is natural to advocate the virtues of policy initiatives in the beginning to gain support and momentum, but such efforts will need to be maintained to reemphasise the benefits of the research and proactively address any false impressions held by the public and other stakeholders, especially as existing evidence suggests that there is still a good deal of public scepticism about clinical evidence of any sort (Carman et al. 2010; Gerber et al. 2010). Besides better public relations on CER, it would prove helpful to design CER studies so that the research addresses questions most important to patients, providers, and policy makers, and to tailor communication of the findings to the needs of different end-users. Coupled with increased stakeholder involvement in the design and conduct of studies, this may help ensure greater understanding of CER, use of the resulting evidence in policy and practice, and commitment to its sustainability. From an implementation science perspective, there is a dearth of research on the use of comparative effectiveness information. In parallel with the scientific aspects of CER, such as research on improved methods, more research is therefore needed on potential policies and practices to encourage the uptake of comparative effectiveness information.

In addition, as previously illustrated, most past and recent arguments against the use of the CER have centred on research affiliated with the federal government. It may be that the private sector therefore has an important role in taking CER forward, especially in terms of using comparative effectiveness studies to support reimbursement and pricing decisions. Although most private insurers have not relied on CER to inform their policies to date, there is some activity in the sector. For example, Blue Cross Blue Shield and Kaiser Permanente both rely on comparative effectiveness information to inform coverage decisions, and a variety of independent research organisations provide evidence reports to insurers, hospitals, and patients. However, increased funding and use of CER in the private sector will require sufficient incentives to enhance the demand for such research and encourage patients and providers to consider effectiveness and costs of care in treatment decisions.

A second significant possible challenge ahead for PCORI and the larger CER initiative relates to the lack of realistic aims and expectations that have been already set. During the debate over the ACA, and in the time since its adoption, proponents of CER have touted its potential to save nearly \$800 billion—an important driver in “bending the cost curve.” However, given the limited power of PCORI and the restrictions placed on Medicare to formally consider CER and, in particular, costs or cost-effectiveness, in its policy making, it is difficult to accept this as a reasonable projection. Available evidence estimating CER’s impact on health expenditures suggest a notably modest reduction in federal spending (< 1 percent) to a potential increase in costs (Basu and Philipson 2010; CBO 2007).

Even countries that formally employ CER and explicitly consider costs and cost-effectiveness in policy making have not realised cost savings. In the UK, for example, while NICE guidance has likely increased the average cost-effectiveness of treatments covered by the NHS, there is no evidence that total spending or the rate of cost increases have been reduced. Furthermore, medical practice is notoriously slow to change in response to new research findings, even in cases where significant cost savings and better patient outcomes are possible. One analysis found that the lag between the discovery of more effective forms of treatment and their incorporation into routine patient care averages seventeen years (Balas and Boren 2000). Other corroborating evidence suggests that practitioners often ignore the findings of large comparative studies (Avorn and Fisher 2010; Maio and Gagne 2010).

There is some potential that the results of CER studies will influence clinical practice, because there is growing pressure from employers and the government to reduce spending, but any reductions will likely not result from CER directly. Therefore, maintaining unattainable expectations of CER’s impact may ultimately result in history revisited when there is a shift in the political tide or PCORI (or CER more generally) falls under the scrutiny of opponents. To guard against future losses of political support and funding, it may be advisable to position and integrate CER as part of a wider range of reform initiatives to improve the quality and value of US health care. For example, CER could be linked to other key health care reforms, such as those related to health information technology and physician payment reforms,

especially those initiatives that have received bipartisan support. This may protect against opponents singling out CER for scrutiny if the aims of the ACA are not achieved. Regardless of the potential political benefits, CER should indeed play a key role in these initiatives to ensure that the evidence generated is linked to other health care reform programs and policies. In addition, although the ACA prohibited PCORI from using CER to dictate policy and practice, it did not bar private payers from using the research to inform their policies. In the private sector, evidence generated by CER studies could feed into value-based insurance designs, where co-pays are reduced for high-value services and increased for low-value services, with the aim to steer treatment choices toward more effective and efficient services (Pearson and Bach 2010). Attaining any cost savings from these options, however, will depend on the extent to which both research findings and treatment choices favour less expensive alternatives, and—again—the willingness of policy makers, physicians, and patients to take account of the relative cost and cost-effectiveness of treatment options in decision making.

Conclusion

Comparative effectiveness research offers significant promise to provide better information on available health care interventions and methods of care delivery and thus ultimately to improve health care decision making and practice. However, as evidenced by previous related experiences that sought to meet similar aims, it is not without challenges. While much has already been done to support an integral and sustained role for CER in the American health care system, considerable efforts will be required in the days ahead to ensure the success and sustainability of PCORI and the overall CER effort.

Conclusions of the thesis

Summary of key thesis findings

The purpose of this thesis was to examine the extent to which good health technology regulation is achieved and the effectiveness of the policy measures regulations (and others) employ to meet the aims or criteria of ‘good regulation’. To accomplish these objectives, the thesis developed and considered a conceptual framework of good regulation and explored its various dimensions in eight studies. Taken together, the studies provide an analysis of:

- The role, processes, policies, and performance of the regulators of the market authorisation and coverage and reimbursement of pharmaceuticals and medical devices in Europe and the US;
- The role and use of technology assessment in health technology regulation and its impact on attaining good regulation; and,
- The factors that impact regulatory performance.

As previously outlined, good health technology regulation can be considered to be attained if regulatory actions are legitimate, accountable, transparent, inclusive to relevant stakeholders, informed by expert and credible advice, effective, and cost-efficient. Such goals are achieved through a set of tools, activities, and discourse through which regulators address their policy objectives and reform priorities. One of the main ways regulators strive to achieve these objectives is to review the available evidence on the technology’s benefits and risks or costs to ensure that available technologies are safe and efficacious and they provide value to patients and the health care system in return for investments made toward their financing.

The following section examines the main thesis findings based on the analyses presented in the eight studies. It then discusses the ways in which health technology regulators can advance existing regulatory practices to attain or maintain good

regulation. Finally, it outlines the limitations of the thesis and areas for further research.

As measured against the ‘good regulation’ framework, the following section explores the extent to which health technology regulators are meeting the outlined criteria; what factors, if any, inhibit or facilitate good regulation; and, the extent to which evidence-based regulatory process aid regulators ability to achieve and maintain the various criteria of good regulation.

Mandate

There is considerable variation in health technology regulators effectiveness in terms of meeting their mandates. In particular, the EMA has attained, over time, a more balanced approach to meeting both of the main tenants of its mandate – industrial support and public health protection. To be sure, industrial objectives remain central to the EMA; however, in recent years, there has been a shift toward greater attention and activities focused on its public health responsibilities. Such improvements include the provision of more reliable and objective information on new medicines for patients (e.g. packaging and leaflet labelling), allowance of conditional marketing authorisations, increased regulation of and funding for post-market data collection and pharmacovigilance. In addition, the agency has increased its interactions and collaboration with other leading medicines agencies, such as the FDA, to better align regulatory processes and harness surveillance activities to improve patient safety. These changes may be attributable to the shift in oversight from DG Enterprise to DG Sanco, expanded responsibilities of the agency, and increased pressure from stakeholder groups to protect public health given the growing number and complexity of new medicines.

Compatively, it is realtively unclear whether regulators involved in device market approval, most notably in Europe, align with their respective mandates. Device authorisation in Europe is highly decentralised, which has introduced barriers to the extent to which Competent Authorities can effectively monitor the work of the Notified Bodies and coordinate robust post-market surveillance systems to monitor and safeguard patient safety. In addition, the diversity in requirements and evidence

standards upheld by the various Notified Bodies hinders their ability to ensure device performance and safety. In both jurisdictions, some moderate- and high-risk devices are approved based on the evidence generated for similar products already on the market and robust post-market surveillance and adverse event reporting of adverse events are limited. Approvals based on less rigorous proof of effectiveness and safety may have a greater chance of later-identified adverse events. Moreover, assessments are conducted by manufacturers and Notified Bodies, which may introduce important conflicts of interest, particularly regarding the attention paid to the balance between effectiveness and risk of safety concerns.

Regulators in Europe and the US recently rejected significant changes to their existing regulatory frameworks that would better ensure the safety and effectiveness of moderate- to high-risk devices before they are marketed to providers and patients. In both cases, the medical device industry vehemently contested such changes, arguing that they would negatively impact patient access and hamper the broader innovation process. Rather, both jurisdictions are instituting or considering incremental measures to improve the systems. Such measures are predominately focused on improving post-market regulation through implementation of the UDI and heightened pharmacovigilance. The more politically controversial (and difficult) changes related to regulatory authority over device approvals and increasing the evidence requirements and robustness of reviews for moderate- to high-risk devices appear to be sidelined at present.

Differences in culture, history, politics, and the organisation and financing of care all have important influences on the mandates of those bodies involved in health technology coverage and reimbursement, especially with regards to the underlying rationale for undertaking HTA to support such decisions. All of the countries examined in the thesis, with the exception of the US, have adopted technology assessment, as part of or the entirety of, their mandate to ensure that funded technologies offer patient benefit at reasonable cost. However, fulfilling their mandates has manifested differently across countries. For example, while all countries aim to ensure that covered technologies benefit patients, they diverge in terms of the importance of value for money, equity, and broader ethical and social considerations. That said, based on the thesis findings, we can see a shift in the

mandate of coverage and reimbursement regulators to ensure funded technologies provide added value to not only patients, but also to other stakeholders and the broader health care system. The mandates of these entities also depend on where they are situated within the health care system; some are integrated into other government authorities, while others are at arms-length to government. It can be argued that in most cases, those entities at arms-length have clearer mandates and tend to be more transparent in not only their role and aims, but also their processes and decisions, which has important (positive) implications for garnering public trust and legitimacy. A good case example is NICE in England. Although there are, of course, exceptions, in general being at arms-length can also help protect regulators from day-to-day politics and safeguard their exercise of specific competencies (expertise) to meet said mandate.

Accountability

In general, health technology regulators have increasingly enhanced and maintained their level of accountability in recent years. These advancements are due, in part, to improvements made to attain some of the criteria, such as due process. Both market authorisation and coverage and reimbursement bodies provide avenues for communicating and justifying their decisions, such as publishing their decision processes and criteria on-line (especially with regards to coverage and reimbursement entities) and offering tailored guidance or other mediums (e.g. EPARs, SMPCs) that outline key evidence considerations and other aspects informing decisions to different end-users (e.g. patients, health professionals). This opens up the opportunity for external scrutiny of the regulators' activities and decisions. NICE, for example, releases new draft guidance and methodological standards for stakeholder comment prior to final release and it undergoes an annual "performance review" by Parliament to ensure it remains "fit for purpose" and provides value to the NHS. Ensuring an appropriate locus of regulatory oversight has also been important for accountability. The EMA, for example, gained greater accountability by way of transferring oversight from DG Enterprise to DG Sanco. In Europe, improvements to the accountability of device regulators are in progress. The Competent Authorities will exercise additional oversight and quality assurance of the

Notified Bodies, and the European Commission will now be involved in selecting and vetting Notified Bodies and maintaining the Eudamed database.

However, challenges to accountability remain. Regulators involved in market authorisation require significant “user fees” from industry to review their products, which may beholden regulators to industry (and therefore facilitate regulatory capture) given that such fees comprise a high proportion of their budgets. Nevertheless, the regulation of health technologies is a high resource activity and is becoming increasingly so with the growing number and complexity of therapies. Such fees not only help support individual product reviews, but also contribute to retaining experts, developing post-market activities, and ensuring that approval time targets are met. Instituting effective due process mechanisms and ensuring proper conflict of interest policies can attenuate the potentials risks to accountability associated with user fees.

Another challenge centres on the degree of regulatory decentralisation, particularly with regards to coverage and reimbursement policy. In those countries with more decentralised health care systems (e.g. Italy) or processes for coverage and reimbursement determinations or implementation of decisions (e.g. Sweden), the lines of accountability are less clear, which can raise concerns about duplication of efforts and authority. This may stymie effective coordination, consistency of decisions, transparency, opportunities for appropriate external performance evaluation, and, ultimately, effectiveness and cost-efficiency.

Due Process

Health technology regulators have increasingly instituted practices and policies to improve due process, especially with regards to the market authorisation and coverage and reimbursement of drugs. The EMA involves patients, health professionals, and manufacturers in its processes and decision making, provides information on new drugs and the regulatory process itself on its website and via other methods, and is working toward greater transparency and open exchange of clinical trial results and the evidence used to support approval determinations. Policies and procedures to maintain due process are less evident in the case of

devices, particularly in Europe. In the US, the FDA places proprietary limits with public reporting of pre-market review of approved devices, recalls, and adverse events, lending public access to evidence on new devices. Moreover, a variety of stakeholders – from regulatory and scientific experts to consumer groups – are involved in the FDA’s work on devices. It is uncertain the extent to which relevant stakeholders are involved in the European system for device approval, given the breadth of Notified Bodies and Competent Authorities involved and a lack of uniform standards or policies in this regard. Such limitations also apply to appropriate levels of open communication about the procedures and decisions adopted by these different entities. In both jurisdictions, however, information on approved (and rejected) devices should be more openly communicated and accessible.

Achievement of due process has proven more central to coverage and reimbursement decisions. Coverage and reimbursement/HTA bodies generally have various opportunities for stakeholder involvement in different areas of the coverage and reimbursement process and clear procedures for their engagement. For instance, many of these bodies (e.g. NICE, IQWiG) have established both patient and industry groups to guide regulatory priorities, evidence and methodological standards, and decision implementation strategies. Manufacturers are increasingly collaborating with these entities to guide research and development efforts and enhance the cost-efficiency and predictability of the coverage and reimbursement process. The US, in particular, has prioritised the patient perspective and experience as central to technology assessment and, to some degree, coverage and reimbursement decisions. While the focus on “patient-centeredness” in the US, notably around the investment in CER and PCORI, may have initially served to garner political acceptance, patients have meaningfully participated in setting the research agenda, standards, and processes for communicating the resulting research. If CER data eventually plays a more significant role in the coverage and reimbursement decisions of CMS and/or commercial payers, the central involvement of patients (and other stakeholders) may lead to greater acceptance of evidence-based payment policies.

Coverage and reimbursement regulators have moved toward greater transparency and public availability of information about their activities, decision

processes, decisions, and modifications or reforms to their governance and policies, with some countries more successful in meeting these ends more than others. NICE in England is arguably the most transparent and inclusive with regards to offering a breadth of publicly-available and accessible information. The institute publishes details about its processes and coverage decisions on its website and tailors guidance to specific users, such as health professionals and patients. Other regulators have increasingly implemented similar approaches, as greater stakeholder involvement and interest in health technology regulation have resulted in growing demand for more accessible and user-friendly information.

Greater openness, transparency, and inclusive participation have likely contributed to a blurring of boundaries between state-centred and self-regulatory (public and private) forms of health technology policy governance, where the regulated (industry) are involved themselves in the regulatory process and setting standards and policies; third parties (e.g. patient, the public, health professionals, media) may be engaged in monitoring and (direct or indirect) enforcement; and, regulators serve an enhanced role in information pooling and exchange and coordination across all involved and/or impacted parties. These trends can be viewed in the increased role of manufacturers in determining regulatory priorities, evidence requirements, and post-market data collection and surveillance. Moreover, patients, the public, and media all increasingly exhibit external pressure on regulators to uphold their mandate, be accountable, exercise open and fair processes, and influence key outcomes of interest, such as timely access to safe technologies, affordable and equitable care, health gains, and managed spending. The ascendance of electronic communications and social media has played a significant role in fostering third party monitoring, evaluation, and enforcement of health technology regulators.

Whether these changes have positive implications for good health technology regulation depends on the circumstances and how regulators respond to stakeholder influence. As previously discussed, the term ‘agency or regulatory capture’ is commonly used to describe an industry exerting undue influence over the agency that is regulating its activity and essentially “capturing” the agency for its own ends. There are certainly instances where the industry has heavily influenced health

technologies regulators, probably more so in terms of market authorisation than for coverage and reimbursement; there are, of course, differences across countries, type of technology, and the therapeutic area addressed (e.g. cancer versus a rare disease affecting a small patient population). Regulatory capture can be seen as more of a risk in those bodies or situations lacking sufficient transparency and accountability. For example, industry has probably had undue influence over regulatory requirements for data transparency and exchange, with the former claiming threats to innovation and competition (via ‘commercial in confidence’ concerns). Another instance is industry-sponsored patient organisation activity, where industry exerts its power and influence to organise patient groups to largely promulgate its own interests and objectives.

With regards to coverage and reimbursement, industry, particularly in the US, has been pivotal in guiding the evidence and methodological requirements supported through technology assessment, such as the adoption of QALYs or cost-effectiveness analysis. The impact of industry influence also depends on the participation and representativeness of other affected parties. Indeed, ensuring due process as well as appropriate accountability encompasses meaningful participation of involved stakeholders, where proportional attention is given to all represented viewpoints and concerns. For instance, well-organised and resourced industry groups might manage to generate more effective pressure on the regulator than a group of consumers and patients. As illustrated in HTA systems to support coverage and reimbursement decisions, while patients and the public are frequently involved in providing input to the process, questions remain regarding whether their engagement is meaningful and not mere lip service, in attempts to consider their needs and preferences. The wider range of actors involved in health technology regulation and the resulting increased demands to have their preferences considered might challenge regulators ability to achieve good process across all circumstances, deter clear lines of independence, accountability and impartiality, and hinder regulatory efficiency. This issue will likely be exacerbated by the increasingly complexity of technology and resulting regulation, which will require greater input from and collaboration between regulators, industry, patients, health professionals, among other impacted stakeholders.

However, regulators, industry, and consumers (and other third parties) are becoming increasingly sophisticated in understanding and overseeing the nuances involved in the availability and financing of new health technologies. Most notably in the last several years, manufacturers have increasingly exercised “pre-emptive” self-regulation (Solomon 2010) to react to concerns from policy makers and affected publics by directly engaging in the actual functions and tasks of regulation. Manufacturers, for example, have sought to engage with the EMA, FDA, and coverage and reimbursement bodies to discuss evidence requirements and study designs and outcomes to meet such regulatory obligations. They also have become more directly involved in post-market data collection and ongoing assessments of technology safety and value. At the same time, we might consider an increase in “industry capture”, where regulators enlist industry to perform regulatory functions. One example is working with industry to advise on developing studies that would meet regulatory requirements, which serves not only to help industry get their products to market, but renders the regulatory decision process more efficient. Regulators can take advantage of this kind of interaction to avoid self-interested pre-emptive self-regulation that serves to remove or delay regulatory issues and developments off the policy agenda.

The blurring of regulatory roles, modes of regulation, and functions may actually aid regulatory effectiveness, particularly with regards to institutional/regulatory adaptation. Health technology regulators can create new regulatory regimes or processes with the flexibility to adjust to problems as they arise and with the necessary buy-in from the private sector (and other external affected parties, such as patient groups and medical associations) to encourage cooperation. Indeed, new regulatory features may emerge from private ordering. If the goal is continuous improvement in achieving regulatory goals, then relevant stakeholders should be working together to better achieve such aims, whereby regulatory activities, processes, methods, and requirements are frequently revisited. This process should be viewed as dynamic; modifications to health technology regulation will certainly be necessitated to balance stakeholder interests and perspectives and appropriately respond to technological advances and changes in the broader health care environment.

Expertise and Impartiality

Both areas of health technology regulation – market authorisation and coverage and reimbursement – rely heavily on the use of experts to inform and guide regulatory processes and decisions. A range of expert committees and working groups are involved as well as individual experts who act as advisors or consultants. Maintaining an expansive network of experts is particularly important in the case of health technology regulation given the breadth and growing complexity and personalisation of new technologies. The primary concern is protecting against potential conflicts of interest that may influence the work and advice of involved experts in ways disconnected from the evidence and standards associated with regulatory decisions or actions. However, the real challenge is discerning when expertise lends to partiality, as different evidence could be selected and/or interpreted differently with similar levels of rationality and with divergent impacts on good regulation.

Based on the thesis findings, impartiality and credibility can be fostered when nomination and selection of experts extends beyond the internal circles of the regulator to allow involvement of experts external to the “establishment”. In addition, some type of peer-review system should be used to safeguard the accuracy and validity of opinions reached by the committee. Conflicts of interest for all involved experts should be collected, monitored, and communicated in a transparency way. Such requirements should extend to external consultants and advisors. At minimum, mechanisms should be established to note and publish all involved experts and their particular role in the regulatory process.

Effectiveness

As noted earlier, regulatory effectiveness can best be understood as whether a regulator delivers intended results or policy objectives. For regulation around market authorisation, this can, in part, be captured by their ability to protect public health. In general, the US system – where FDA has a mandate to regulate both pharmaceuticals and devices – may be more conducive to achieving greater oversight, coordination, pool of expertise, and transparency, and therefore may be more effective in

protecting public health. The FDA's mandate also guides and reinforces its principal orientation toward public health protection. In Europe, the priority of public health in regulation is less clear. The EMA has always been part of a broader strategy to support industry and economic growth within the EU and the main bodies that approve devices are private, third party entities with commercial interests in authorising devices. One argument put forth against the US system of market authorisation is that it is overly risk adverse, slow, and cumbersome, resulting in protracted availability of beneficial, new technologies to better public health. Although this may have been true in the past, recent evidence suggests that approval times for drugs are largely equivalent between the US and Europe (with the FDA perhaps somewhat faster than the EMA). For devices, the European CE marking process is faster than the 510(k) or PMA process in the US – anecdotal information from Notified Bodies suggests that the process takes one to three months, excluding sponsor time, while the FDA takes an average of eight to 13 months to review an application (FDA 2012). However, it takes considerably longer in Europe to gain coverage and reimbursement for a device (and drugs) than in the US. Certainly speed of approval is only one criteria of effective performance toward public health protection and must be considered against policies and practices to ensure a robust and thorough assessment of a technology's risks, benefits, and costs. Nevertheless, health technology regulators have been pivotal in bringing innovations with significant public health impact to market.

As evidenced in the thesis, there are key differences between the effectiveness of existing regulation for drugs and devices. Overall, pharmaceutical regulators require more robust evidence of safety, efficacy, and effectiveness (and, where applied, cost-effectiveness) and operate more extensive systems of post-market surveillance than expected for devices. In both jurisdictions, some moderate- and high-risk devices are approved based on the evidence generated for similar products already on the market, reporting of adverse events remains voluntary for certain users, and information on approved (and rejected) devices is insufficiently shared with affected parties. In Europe, devices are marketed with less rigorous proof of effectiveness and may have a greater chance of later-identified adverse events. Moreover, assessments are conducted by manufacturers and Notified Bodies, which may introduce important conflicts of interest, particularly regarding the attention

paid to the balance between effectiveness and risk of safety concerns. Conversely, however, the systems may offer regulatory flexibility to account for the diverse range of devices on the market.

In general, the effectiveness of coverage and reimbursement regulation has improved with the growing sophistication of these processes, mostly notably in Europe. Governments apply several strategies in parallel to ensure national policy objectives are met effectively within the constraints of their respective systems, with increased focus on evidence-based approaches to ensure value for money and improve quality of care. While these approaches have challenges and limitations as discussed, there is some evidence that they have led to lower prices, better patient affordability, reduced patient and payer expenditures, and improved efficiency. Moreover, in almost all European countries, positive lists (formularies) are fairly comprehensive and cost-sharing is usually low, allowing patient access to medically-needed therapies. If a technology is not publicly covered, it may be available through (voluntary) private health insurance, although individuals may be more likely to pay out-of-pocket for unlisted therapies.

The US operates a decentralised system of coverage and reimbursement and has generally taken a less cost-conscious approach to technology adoption, access, and financing. At the Federal level, Medicare is prohibited from negotiating drug prices and explicitly considering costs or cost-effectiveness in making coverage determinations, both of which have proved politically contentious. Moreover, very few technologies actually undergo a formal national coverage determination (NCD). Private payers employ mechanisms to manage access to and efficiency of health technologies, but these tend to be focused on the demand side through utilisation management, patient cost-sharing, and physician payment and delivery reforms. A minority of private payers use formal HTA processes or other supply side strategies. Similar to Medicare, when evidence of value is considered, private payers tend to consider effectiveness, not costs or cost-effectiveness (at least not explicitly). Compared to Europe, the US pays more for technologies and health care services, and access to care is restricted based upon health insurance coverage and ability to pay. Furthermore, the high spending levels on health care have not produced commensurate health outcome gains. That said, as a result of recent health care

reforms and a push to achieve the triple aims in US health care (improve population health, enhance patients' experience of care, and reduce costs), payers have considered and implemented some strategies focused on value, such as value-based insurance design.

In particular to health technology assessment, its impact on effective regulation depends on a range of factors. As noted in the discussion on mandate, technology assessment has probably been most effective in achieving and maintaining good regulation in countries where the national HTA bodies/payers have the regulatory and legislative authority or mandate to assess technologies and use the resulting evidence in decision making, and have a variety of policy instruments available to implement such decisions. These features also have a positive effect on securing legitimacy and accountability.

Another important factor centres on data availability and methodological considerations. As demonstrated, the available evidence from which to base regulatory decisions is typically better for drugs than devices; lacks sufficient comparison to existing therapies; is often limited for therapies that address a specialised subgroup of patients; and, inadequately reflects “real world” use of the technology in practice. Meanwhile, the methods used in technology assessments have different advantages and disadvantages that have both positive and negative implications for regulatory effectiveness. Moreover, adopted methods often align with different national perspectives and societal norms as to what constitutes effective health technology regulation. So what may work and considered viable in one country may differ from others. A good example is the different use of value metrics (e.g., cost per QALY, efficiency frontier, added value ranking) across countries discussed earlier in the thesis. All of these issues have added complexity to both assessments and the overall regulatory process, given that many of these concerns must be balanced adequately if effectiveness regulation is to be realized. Such complexity will only grow as more diverse and specialised technologies enter the market. A number of actions could advance progress on some of these issues. For instance, RCTs are involved and costly and are not necessarily “fit for purpose” in all contexts or for all technologies, especially given the drive internationally toward the generation and use of real world effectiveness data. Regulators should consider other

study designs and methods, such as pragmatic clinical trials, observational studies, meta-analyses, among others to ascertain the benefits and costs of health technologies⁹². Recent increased investments in electronic health records, registries, large post-market data networks, and post-market surveillance systems should help facilitate the adoption and usability of these alternative methods. Another option, which has been explored throughout the thesis, is the promulgation of post-market solutions, such as CED and risk-sharing agreements. While these approaches hold promise and have piqued the interest of payers and manufactures alike, their lack of transparency, predictability, funding, and methodological and procedural development has hampered realising their full benefit to date. As a result, payers in both the US and Europe have moved toward use of rebates and discounts (essentially financial contracting) due to their relative ease and timeliness compared to outcomes-based agreements. However, continued interest in these arrangements and establishment of a more robust post-market data collection and monitoring infrastructure will likely have positive effects on the ability of the regulatory system to adequately capture and use downstream evidence on health technologies. In concert, incentives will need to be implemented to encourage data collection, use of evidence in downstream decision making, and any resulting changes in policy and practice.

Additionally, the impact of HTA on regulatory effectiveness depends on the particular circumstances surrounding an assessment or regulatory decision. Such circumstances are more qualitative in nature and typically have indirect attenuating effects on regulatory outcomes. For instance, decision makers may not fully understand or accept an assessment, resulting evidence, or final recommendation, and even if they do, the consequences (e.g. a negative recommendation to adopt a given technology) may not always be accepted, especially if it results in challenge from industry, clinicians, patients and the general public. To that end, as evidenced throughout the thesis, assessments may not sufficiently capture the priorities, preferences, and values of decision makers, providers, patients, and broader society, which would hinder perceptions of due process and legitimacy. Consequently, coverage and reimbursement decisions may deviate from the available evidence. This may or may not be problematic, depending on the particular context or

⁹² See Sorenson et al. (2011b) for discussion of the advantages and limitations of these approaches.

circumstances. Brown and Gusmano (2013) emphasise the importance of recognising that the goal of technology assessment should not be to eliminate but to enrich political deliberations that govern what societies fund and receive from their respective health systems. To that end, the political climate is influential. For example, HTA has not received sustained political and stakeholder support in the US, which has stymied its use to facilitate regulatory effectiveness. A number of case examples explored in the thesis highlight these challenges, including PCORI and previous agencies, such as Oregon Medicaid, and valuing advanced cancer care. The (failed) Oregon Medicaid experiment highlights how evidence/methodological complexity and many of the aforementioned circumstantial, social, and political factors can derail regulatory efforts to meet coverage and reimbursement aims. While those involved used a range of viable methods to rank the cost-effectiveness of health services to determine coverage by the Oregon Health Plan, the approach instigated national upheaval due to misunderstanding and disagreement of the ranking process, ethical and legal concerns, and political tensions around rationing and placing monetary value on medical conditions.

Overall, however, it is difficult to ascertain whether health technology assessment has led to more effective regulatory actions and decisions, as the counterfactual is unknown (i.e. what decision would transpire in the absence of technology assessment) and the evidence on outcomes is limited. Nevertheless, on balance it has probably resulted in regulation (notably with regards to coverage and reimbursement) that is more transparent, inclusive, predictable, consistent, and, ultimately, effective. Of course, attaining these endpoints has been – and still remains – a process of trial-and-error and continuous learning and evolution for regulators. An important consideration moving forward is whether wider expectations of HTA (and health technology regulators in general) are realistic and attainable. In other words, are we expecting too much from health technology regulation? There are limits on resources, time, and the speed and breadth of adaptation that can be expected of regulators, especially considering the quickly evolving and complex health technology landscape. On balance, health technology regulators have been responsive to ongoing shifts in the marketplace. They have introduced new methods and processes to address evolving needs and challenges and have engaged with other public and private actors to expand and complement their

expertise and resources in efforts to ensure good regulation. Regulators will need to continue to be responsive in order to effectively address an increasingly complex policy area.

Cost-Efficiency

Ascertaining the cost-efficiency of health technology regulation is the most challenging to evaluate, given its links to the other criteria. The establishment of health technology regulatory agencies, such as the EMA and some HTA bodies, have undoubtedly saved member states time and effort in evaluating new technology and, in the case of the EMA, has created a homogenous market authorisation policy for drugs throughout the EU. Time to market authorisation and/or coverage and reimbursement approval may be one marker for cost-efficiency, but as discussed in study 1, factors external to the agency might impact productive efficiency, such as variations in timing of product entry into the different markets. Factors internal to the regulatory body may also influence efficiency, such as the robustness of technology reviews, stakeholder involvement, and mechanisms to ensure transparency. Although potentially detrimental to efficiency, these factors are generally viewed as essential to attaining the other criteria of good regulation. The right balance between efficiency and some of these other regulatory aims will likely differ across affected actors. Industry, for example, considers protracted approval and coverage and reimbursement processes problematic; others, such as the public, may equate slower timelines as necessary to protect public health and ensure new technologies provide good value for money. Predictability of the regulatory process also lends to cost-efficiency. Standards, clear evidence requirements, industry guidance, and other mechanisms all support enhanced predictability and protect against duplication and unnecessary cost. Health technology regulators have generally instituted such approaches to meet these aims (in parallel to enhancing transparency).

Cost-efficiency is a central objective for coverage and reimbursement regulators. In particular, coverage and reimbursement bodies, primarily through the use of HTA, want to ensure that public (or private) public investments are used to maximise health gain. As discussed previously, the use of QALYs and cost per QALY is one method employed by countries to ensure decisions are allocatively

efficient⁹³. Under these circumstances, a regulator would not be efficient in this regard if it covered therapies with low levels of cost-effectiveness (or other measures of efficiency), as limited funds could be better deployed for therapies offering higher potential health gain. As coverage and reimbursement bodies have adopted methods to ensure funds are allocated toward those technologies that provide value (and, in particular, value for money) to the health care system, greater cost-efficiencies have been achieved.

However, claims to cost-efficiency are more complex to ascertain. In most cases, coverage and reimbursement bodies review only select health technologies. This has several implications. First, those technologies deemed cost-effective may displace equally (or more) cost-effective technologies already on the market that have not undergone review. The fact that most cost-effective technologies are also cost-increasing contributes to this issue, as local purchasers of health care may be required to remove existing health care services in order to fund the recommended new technology. Unless local purchasers receive guidance on which services should be prioritised within budget constraints, they may displace different services (introducing differential geographic access to care) or, as mentioned, scale back on services that provide even greater value. Second, existing or new technologies that are not selected for review, yet covered, may be obsolete and/or of low value, which not only results in care inefficiencies, but also consumes limited resources that could be directed toward more cost-efficient therapies.

Moreover, an important trade-off to cost-efficiency, especially with regards to cost-effectiveness analysis, is the fact that it is concerned principally – as traditionally considered – with the total increase in health gain generated by a health technology, but not with how that health gain is distributed, leading to distributional or equity concerns. Such concerns generally fall into two perspectives – one being that the outcomes of regulation are equitable if they derive from a fair process, with the second focused on the fairness of the outcomes themselves (although the ethical acceptability of some processes may come in question). One issue encapsulated in both perspectives is the need to consider social values or perspectives in judgments

⁹³ Note that QALYs are just one method used to meet efficiency targets. Other countries, such as Germany and France, employ other approaches, as discussed in studies 4 and 5.

about fair access to health technologies and the equitable distribution of health outcomes amongst the population. Attendance to equity has many challenges, ranging from measurement complexity to deriving consensus on how value-laden assumptions should be incorporated into assessments (particularly the quantification of the comparative benefits and costs of technologies), which may increase the costs of regulation.

Recent developments in health technology regulation have had likely mixed effects on efficiency. The increased number and complexity of health technologies has resulted in an expansion in the scope and complexity of regulators' responsibilities. For instance, the centralised procedure for drug market authorisation in Europe now extends to orphan drugs, HIV/AIDS, cancer, diabetes, and mental health therapies, as well as to generics, biosimilars, and non-prescription medicines. Regulators are also increasingly required to oversee and participate in post-market activities. Without a commensurate rise in resources and funding, such changes have probably hampered efficiency. The movement towards collaboration with other regulatory agencies (e.g. EMA and FDA, EMA and HTA bodies in Europe), international experts, and industry has increased efficiency on some level, but the resources and time required for effective coordination may lessen such gains. This is also true of the introduction of conditional approval and reimbursement strategies (e.g. CED and risk-sharing agreements), where (certain subpopulations of) patients may be granted access to potentially beneficial technology on the condition additional data is collected post-market to substantiate the regulatory decision. Cost-efficiency gains will only be realised if those post-market efforts are efficient and effectively address the outstanding uncertainty regarding the technology. Relevant studies in the thesis demonstrated that, to date, these approaches have not produced sufficient efficiency gains for the cost. Greater post-market data collection and assessment should also be used to support disinvestment of obsolete, low value, and unsafe technologies. Some countries have made progress to this end, but formal processes for identifying topics for disinvestment and removing them from practice (or deterring use) once identified as low value are required if disinvestment is to be effectively adopted.

In addition, the drive for greater efficiency requires not only assessing value, but also the creation and adoption of incentives that ensure that users effectively translate such evidence into practice, as previously discussed. As evidenced in the thesis, regulators have increasingly recognised and addressed this issue through different strategies to enhance evidence dissemination and translation. PCORI, for example, has a dedicated program and funding to ensure CER results are integrated into current practice. As more and more evidence is produced on the effectiveness and cost-effectiveness of health care technologies (and other services), issues around translation and implementation will be a priority for regulators in the years ahead.

In attempts to address the efficiency-equity trade-off, regulators, particularly in Europe, have strived to better incorporate societal values and equity concerns into technology assessments through citizen councils or working groups that discuss key social and equity concerns and priorities, broad stakeholder involvement opportunities, transparency of process, and improved value assessment methods to reflect social preferences and priorities, many of which were discussed throughout the thesis. The impact of these strategies on balancing both cost-efficiency and equity concerns remains to be substantiated.

This thesis posited that meeting the criteria of ‘good’ health technology regulation is complex and challenging, because of the inherent uncertainty regarding the benefits and risks of new technologies, their growing diversity and complexity, the limitations of existing study designs and assessment methods, and the increased demands placed on regulators to meet sometimes conflicting and continuously evolving objectives and expectations. The impact of evidence-based approaches and policies to improve health technology regulation in particular depends on a number of factors, such as what outcomes are assessed and the quality of the evidence available; the strength of the link between the evidence reviewed or generated and decision making; stakeholder preferences and interests and whether these are taken into account and, if so, how; the organisation and traditions of the institutions involved as well as the broader health care system; and, the extent to which all of these actions transpire with sufficient accountability, due process, expertise and impartiality, and efficiency.

The thesis demonstrated that procedural dimensions of good regulation, such as transparency, stakeholder involvement, accountability, and impartiality are closely linked to achieving effective and efficient regulation. Furthermore, it showed the challenge in adhering to these principles in all circumstances and contexts, and that achieving one principle may necessitate trade-offs in maintaining another. For example, inclusive stakeholder involvement in the HTA process for coverage and reimbursement may slow the decision process (weakening efficiency) and heighten the risk of potential conflict of interest or regulatory capture. Regulators have made progress on certain principles, namely stakeholder involvement and, to some degree, accountability and legitimacy. However, additional improvements are needed to improve the impartiality and transparency of health technology regulation. In particular, strong conflict-of-interest and public information sharing policies are needed. On balance, however, regulators continue to strive for good regulation, although to varying degrees of success across jurisdictions and type of technology. Like most of health care, movement toward achieving the various criteria of good regulation will likely be incremental, especially considering the often step-wise (and additive) pattern of technological innovation.

One of the main propositions of the thesis related to the priority of the different criteria outlined in the good regulation framework, namely outcomes (effectiveness, cost-efficiency) compared to the procedural criteria (mandate, accountability, due process, expertise and impartiality). In particular, I proposed that while good process in and of itself is not sufficient to achieve and maintain good regulation, it is central to this aim in a couple of ways. First, as intimated previously, attaining the procedural criteria set out in the framework facilitates the likelihood of an effective regulatory outcome(s). If the EMA, for example, exercises a clear mandate, provides sufficient justification to its processes and decisions, involves a balanced representation of stakeholders and experts, and openly communicates relevant information (e.g. processes and criteria for decisions, changes in procedure, activities and priorities), it is more likely that its decisions will be effective and, perhaps, cost-efficient. Second, good process helps maintain the credibility and sustainability of regulatory decisions or policies if and when they come under question. As previously noted, this is particularly important in the case of health technology regulation, which, as previously noted, is increasingly complex, often under external scrutiny,

and involves a broad range of affected stakeholders with often times competing interests. For example, if a regulator is criticised or challenged by a particular decision, the fact that it was arrived at by way of a transparent, inclusive, accountable, and independent process may protect against undue stakeholder influence and facilitate increased support and adoption. Upholding the procedural criteria seems to be especially important for regulatory decisions on high impact conditions with considerable stakeholder and public attention (and thus potential conflicts), such as cancer.

Given the challenges in adhering to all of the good regulation criteria in all circumstances and contexts, and that achieving one criterion may necessitate trade-offs in maintaining another, it raises the question of whether or how the individual criterion should be ranked. Ultimately, regulators are first and foremost required to fulfil their mandates and deliver effective outcomes. In the case of those regulators overseeing market authorisation (e.g. EMA), for instance, it is vital that their actions and processes protect public health and support the availability of important new therapies. The ranking of the remaining criteria, however, becomes more nebulous and complex, as many of the procedural criteria are interconnected and bolstered by good regulatory practices associated with another criterion. For instance, adequate transparency of process and outcomes will help foster adequate accountability. In addition, cost-efficient regulation requires that all of the other criteria be met. Ranking the importance of the different criteria might be possible and beneficial in terms of assessing how different affected stakeholders view good regulation, the criterion that should be prioritised which criteria are priorities, and whether such perspectives of importance differ across contexts or regulator. This could be an important exercise considering the blurring lines between state-centered and self-regulation, and the importance of stakeholder engagement and acceptance of health technology regulatory processes and decisions.

Policy and practical recommendations to improve good health technology regulation

While the thesis demonstrates that existing health technology regulation, on balance, frequently meets many of the dimensions of good regulation (and increasingly so), there remain limitations and challenges. Based on the thesis findings, I discuss four important overarching mechanisms that regulators and other relevant stakeholders should consider and potentially adopt to further advance health technology regulation.

Align aspects of market authorisation and coverage, reimbursement, and pricing for new technologies

A considerable degree of overlap currently exists between the market authorisation and coverage and reimbursement of health technologies, and harmonisation of certain aspects is both needed and feasible. Greater alignment would improve the efficiency of regulatory processes, enhance the clarity and predictability of regulation, raise evidence standards and resulting evidence quality, and increase the likelihood of bringing health gain to the populations regulators serve. In particular, the scientific/evidence requirements that underpin the evaluation of safety, efficacy, and effectiveness lend themselves to a more centralised, common approach. Decisions on coverage, reimbursement, and pricing often reflect local and political considerations that drive willingness and ability to pay, and therefore uniformity of processes or decisions is not likely feasible or even desirable.

One mechanism to facilitate the effectiveness and cost-efficiency of regulation is joint market authorisation/HTA/coverage scientific advice for industry on the design for pre- and post-market evaluations (especially Phase II/III and Phase IV studies) for specific therapeutic conditions. The joint scientific advice could address such matters as appropriate comparators/hierarchy of evidence (e.g. inclusion and choice of active comparator), meaningful clinical endpoints/surrogates where appropriate, ideal study populations and subgroups, and useful effectiveness measures (and patient-reported outcomes) that can be addressed during clinical development and subsequently used to inform coverage and reimbursement

decisions. The results of a pilot programme of joint advice in Sweden suggests that the approach (and its benefits) will need to be clearly communicated to and discussed with industry, as they are used to two separate processes, have organised their companies (and functions) accordingly, and are cautious of new approaches that could increase the costs of development. One particular benefit for industry from early scientific dialogue with regulators is assistance with aligning their research and development strategies so as to contribute to the maximisation of the social objective of providing the best possible quality and sustainable care for the largest number of patients – an outcome that will be increasingly required in the future as the gap between the costs and pace of technological innovation and economic growth is likely to widen in the future (Drummond et al. 2013).

In addition, pre-regulatory dialogue and collaboration between regulators and other stakeholders, namely clinicians and patients, could help identify evidence gaps about the value of a technology and provide guidance for the design of future trials that address recurring deficiencies. Indeed, systematic reviews intended to inform clinical and health policy decisions routinely conclude that the evidence published from thousands of clinical trials each year is inadequate to make an informed decision about the real-world value of a technology. More and earlier collaboration on trial design could therefore help address this problem. Both of the aforementioned recommendations illustrate instances where “pre-emptive self-regulation” via direct involvement of both regulators and manufacturers may lead to better and more efficient regulatory outcomes.

Another approach entails providing HTA bodies/payers full access to the evidence reviewed for market authorisation or, at minimum, revising the EPARs for drugs (and instituting them in the case devices) to allow better information exchange on their benefits and risks for use in effectiveness assessments. Such revisions should be guided by discussion and information sharing between licensing bodies and HTA bodies/payers to ensue the information needs of all parties are met in the most transparent way.

Finally, to address the low level of evidence available at market launch of a new technology (and when market authorisation and coverage and reimbursement

decisions are made), post-market studies are increasingly required by regulators to substantiate safety and effectiveness. This is therefore one more area of potential alignment, where regulators could develop a coordinated data collection and analysis plan, in collaboration with manufacturers.

Increase pre-market evidence requirements and flexibility on accepted study designs/approaches

In the case of pharmaceuticals, regulators should require evidence of comparative efficacy for most new drugs (one exception would be when no treatment alternative exists in a therapeutic area). Such evidence at the time of authorisation would help ensure that the most beneficial and safe treatments reach patients and that limited health care resources are invested wisely. To that end, comparative efficacy data would serve as an important input into subsequent technology assessment conducted by HTA bodies/payers – data that is often poor or missing. Manufacturers can currently therefore differentiate their products from competitors on factors unrelated to therapeutic value (Stafford et al. 2010). However, as previously discussed, there are challenges associated with existing methods to assess comparative efficacy (e.g. direct, head-to-head RCTs with an active comparator), especially with regards to demonstrating superiority, which places greater demands on the size and complexity of trials (Sorenson et al. 2011b). Greater flexibility to develop and accept alternative study designs and approaches is therefore needed. Sorenson et al. (2011b) outline the potential advantages and disadvantages of RCTs with an active comparator versus alternative designs, including pragmatic clinical trials and network meta-analyses. In general, available study designs should be considered as potential complementary methodological tools. A number of actions could help foster the use of comparative efficacy evidence in drug approvals, including arriving at a consensus regarding appropriate and realistic comparative efficacy evidence requirements and study design standards; establishing an independent expert panel, bringing together regulatory bodies (including HTA agencies) and other relevant entities or investigators involved in designing and overseeing studies, to provide guidance on study particulars (e.g. appropriate comparators, sample size requirements to demonstrate superiority, equivalence or

non-inferiority) and create better linkages between pre- and post-market studies (see above).

While comparative data on devices at the time of market authorisation would be ideal, an important first step is to ensure that moderate- and high-risk devices are approved based on robust evidence that they benefit patients, not only that they perform as intended by the manufacturer. The fact that similar devices are on the market should not negate requirements for the new devices to demonstrate safety and efficacy. In Europe, this would require increasing the evidence requirements for moderate- and high-risk devices and ensuring that well-designed RCTs for high-risk devices, using blinding and hard endpoints whenever possible, are submitted to regulators. The latter point applies to the US as well as developing more stringent standards on acceptable predicates for moderate-risk devices reviewed by way of the 510(k) route.

Enhance adaptive or conditional market approval and coverage policies

An underlying theme running throughout the thesis is that regulators face problems of insufficient evidence on the benefits, risks, and costs of new health technologies at the time of decision making. This issue is set to grow in magnitude with greater research and development efforts focused on targeted therapies and personalised medicine, where patient populations are small and limited treatment alternatives exist. Therefore, generating evidence about new health technologies is and should be increasingly seen as an activity that occurs throughout the entire lifecycle of the technology, rather a single one-off review at product launch. Consequently, the use of adaptive or conditional approval and coverage and reimbursement policies may be increasingly adopted by regulators to allow for step-wise learning about a technology under conditions of uncertainty, with iterative phases of evidence collection and regulatory evaluation. As previously discussed, such approaches could be increasingly viable as regulators, in collaboration with other stakeholders, develop and advance post-market data collection systems (via the use of registries, “big” data networks, meta-analyses, electronic health records and other health technology mediums). Collaborative studies between manufacturers,

industry, and payers may also help lend greater credibility to the results and improved effectiveness and efficiency to downstream value assessments.

Develop and institute mechanisms to encourage the use of evidence in decision making

Given the resources dedicated to generate robust evidence on the benefits and risks of health technologies, it is important that it be used to inform and advance decision making. At present, such evidence has a fairly well understood and consistent role in market authorisation decisions (whether the relevant evidence requirements are sufficient is debatable) and a somewhat uncertain to formal link to coverage decisions (and, in some cases, reimbursement determinations). Therefore, regulators should implement additional mechanisms to ensure that investments in research are subsequently used to enhance decisions on coverage, reimbursement, and pricing across the range of available health technologies. However, to capitalise fully on investments in technology assessment or comparative effectiveness research, use of the research should extend beyond market authorisation and coverage and reimbursement. Sufficient infrastructure and incentives should be developed and implemented to encourage providers, patients, and other consumers to consider available evidence on the relative benefits and costs of a technology or intervention prior to use. Simply making the evidence available is insufficient. Rather, more creative and user-friendly mechanisms are needed, such as integrating evidence in physician decision support tools, physician performance standards and payment, smart phone applications and other health information technology platforms used by patients and consumers, as well as developing medical school and continued medical education curricula that educates new and practicing clinicians on the role and importance of technology assessment and comparative effectiveness research.

Limitations of the thesis

While the thesis provides an in-depth analysis of various aspects of health technology regulation, there are some limitations worth discussion. First, the individual papers possess a few shortcomings. In study 1, it may have been advantageous to comparatively evaluate the FDA in the US and the EMA in Europe together using the framework. However, evaluating the EMA alone was a significant undertaking and therefore beyond the feasibility of the thesis. Also, because the FDA has been thoroughly evaluated previously (Carpenter 2006, 2008, 2010), focusing study 1 on the EMA filled an important research gap.

With regards to studies 2 and 8, a comprehensive array of available published and grey literature (and, where applicable, legislative documentation) were collected and analysed on medical device regulation and past and present CER initiatives, respectively. However, interviewing various thought leaders (e.g. FDA leadership, European Commission staff, and regulation experts in the case of study 2; individuals involved with the various CER agencies and organisations regarding study 8) may have provided different and more nuanced insights not readily apparent based on in-depth analysis of secondary data sources. Moreover, the medical device reforms, particularly in Europe, and CER initiatives in the US are still being finalised and under development. Therefore, some information in the papers may not be entirely reflective of the changes ultimately adopted. Nevertheless, the documents reviewed were published at a time when relevant developments were still unfolding, and therefore served to inform current debates in these areas.

Given that studies 3, 4, and 5 were largely comparative review papers that cover an array of countries, they were not able to go into significant depth on any one country or policy. Therefore, the papers may overlook some meaningful details on the different national approaches. However, the goal of studies 3 and 5, in particular, were to provide a cross-country comparative overview of different pricing and coverage policy mechanisms, with a more detailed analysis of the drug evaluation process and HTA especially. An even further in-depth exploration of HTA in Europe can be found in my book (Sorenson et al. 2008a). The three papers, particularly studies 3 and 5, may also perhaps benefit from including other countries.

In study 6, the opinions and experiences of the experts interviewed may not be reflective of all those involved in and affected by CED policies internationally. However, the experts participating in the study are leading internationally-recognised experts in the areas of health technology assessment, comparative effectiveness research, evidence-based policy making, and coverage and reimbursement policy. Those participants from payer agencies and industry are senior leaders within their respective organisations, and those from payer agencies have direct experience related to the development of their respective national CED policies. Moreover, I strived to involve a variety of key stakeholder groups engaged with CED policies internationally in the study to ensure that a range of perspectives and experiences were captured – scientific, technical/methodological, procedural, and political. In addition, the opinions of the experts interviewed were complemented by and verified against the information gathered from the literature review and, importantly, the interviews were carried out until sufficient saturation in the responses were achieved (both in terms of information on the national CED policies and on the opportunities and challenges of CED). The findings of the study were also presented and verified for accuracy and relevancy at two European meetings before finalising the paper(s), including one workshop on CED in the context of devices (with European policy makers, academics, and device industry representatives in attendance) and one annual conference on pharmacoeconomics and outcomes research. Another limitation worth noting is the incomplete information gathered on France. It proved extremely difficult to obtain sufficient information from the available literature and grey sources and secure interviews with more than one French expert. However, France does not represent a leader/innovator in this area, so it is unlikely that the overall results of the study were impacted.

Finally, with regards to study 7, although the paper attempted to include all available literature in the analysis, the existing evidence on stakeholder perceptions of the value of advanced cancer drugs is limited. As noted in the paper, there are number of studies focused on oncologists' views, but relatively few to no studies examining the perspectives of patients, policy makers, carers, and the general public. Therefore, the evidence reviewed may not adequately capture how these various groups value these therapies.

With regards to the limitations of the overall thesis, the conceptual framework of ‘good regulation’ developed and used to underpin the analysis in the various studies arguably cannot address all of the complexities of health technology regulation. Additionally, the various dimensions or criteria may not be considered universally applicable. However, the thesis strived to address these issues by integrating the academic and practitioner criteria together into the framework. Such criteria are not only applicable to health technology regulation, but the health arena more generally and in other policy contexts (e.g. environmental regulation, food regulation). One important way the framework might be improved is to more finely tune the dimensions. For example, the framework presumes that indicators of target-setting and responsiveness are captured within the six dimensions, but it may be advantageous to include them as separate criteria. Moreover, given the importance of transparency to good regulation, it may be helpful to make it a stand-alone criterion rather than include it within the due process criterion, which also encompasses stakeholder involvement, trust, and other measures of good process in regulation. Other criteria not included in the framework, such as equity, may be warrant inclusion, especially in the context of coverage and reimbursement. Finally, several of the studies (particularly study 8) highlighted the influence of politics on good regulation. As intimated earlier, politics matter – that is, they can have both an enabling and disabling impact on regulatory performance and quality (and, certainly on the sustainability of new regulatory entities). Therefore, it may be of value to add politics as an additional dimension to the ‘good regulation’ framework. Of course, this may be difficult to both define and measure, and arguably it may overlap with some of the other dimensions. For example, sufficient accountability, due process, and expertise and impartiality should, in principle, help protect against inappropriate or deleterious regulatory capture from industry and interest groups. The impact of politics, however, cannot be understated and conceptual framework of good regulation should reflect and capture its potential influence.

The thesis may have also benefited from investigating the same jurisdictions throughout all of the papers. However, the paper focused on the jurisdiction(s) that were most appropriate for the research issue(s) under examination. Furthermore, in situations where the thesis delved into national comparisons or case studies, the

countries were selected based on relevance, availability of information, and research need.

Further research

In spite of growing policy and academic interest in and use of technology assessment, the thesis, particularly study 5, highlighted a dearth of research on resulting outcomes (and thus effectiveness) from the approach, particularly in terms of health care budgets, clinical practice, and health outcomes. Perhaps this is not so surprising given the data challenges of assessing outcomes and their attribution to technology assessment. Beyond the methodological difficulties, though, one may detect a lower priority level accorded to outcomes among governments wholly absorbed with building HTA infrastructure, and simultaneously trying to avoid the political pitfalls related to perception that technology assessment is simply the rationing of health care disguised as rigorous policy analysis. A better understanding of the impact of technology assessment or comparative effectiveness research on outcomes is particularly important to maintain and garner continued support for this approach. In particular to the US context, demonstrating impact is important for long-term adoption and use of such research. That said, research and subsequent analysis on the merits of this approach should be cognisant of the overall limitations of this approach and its intended aims. For instance, in most countries, it was never intended to reduce costs, but that is how it is often touted by supporters or opponents (when expenditures continue to rise).

Another area for further research centres on the methods for assessing the value of health technologies. Commentators have noted problems with the QALY as a measure of the social value of an intervention (Baker et al. 2010; Dolan et al. 2008). Other alternatives to QALYs have been proposed such as the efficiency frontier approach adopted by the IQWiG in Germany, estimating willingness to pay through contingent valuation, and discrete choice experiments (Bridges et al. 2010; Neumann et al. 2012). To date, however, there is not enough research on these approaches to provide an assessment their feasibility, usefulness, and acceptability by decision makers. Further evaluation of these methods should also consider their respective advantages and disadvantages for different types of technologies. Devices

and drugs differ in important ways that impact methodological/assessment considerations (Drummond et al. 2009; Sorenson et al. 2011a).

In order for evidence on the comparative risks and benefits of new health technologies to improve decision making and, ultimately, patient and public health, it must be made available in a timely and understandable way and used by intended end-users. Considerable resources are dedicated to clinical studies, comparative effectiveness research, technology assessments, and the like, but investments in implementation science are limited. It is therefore imperative to better understand the perceptions, needs, and preferences of different stakeholder groups and to identify levers that both facilitate and hinder access and consideration of evidence in decision making.

In addition to further research on the outcomes of technology assessment and CER, we need to better understand the political and social considerations involved in evidence-based regulation or policy making, as previously mentioned. Such research would enable more effective analysis of decision making around health technologies, with respect to the interplay between evidence, political and institutional dynamics, “pressure politics”, stakeholder values and interests, and balancing technical, social, and political priorities. In essence, focused analysis along these lines would allow deep understanding of the impacts of political dynamics and the movement toward greater acceptance and adoption of self-regulation in health technology regulation. In addition, further inquiry is needed to better understand and assess the impact of increasingly blurred regulatory boundaries between state and self-regulation.

Along those lines, considering the growing policy interest in and action toward harmonisation of some aspects of HTA (and even market authorisation and coverage and reimbursement, to some extent) as well as increased collaboration between regulators and payers (e.g. FDA-CMS parallel review for certain devices, Green Park Collaborative, joint scientific advice), it would seem beneficial to evaluate the effectiveness and cost-effectiveness of such arrangements and how they might be improved. Indeed, research in this area is needed to better understand in what ways multi-stakeholder collaboration and the harmonisation of certain processes and/or

evidence requirements improves or hampers effective and efficient health technology regulation.

Finally, while the thesis applied the conceptual framework of good regulation in the context of the health technology regulation, it would be informative to further test and substantiate the framework by employing it in other policy areas. For example, the framework could be applied in full to individual HTA bodies or other health care regulators (e.g. food, health care professionals), in addition to non-health regulators, and across different jurisdictions. Such research would presumably lend greater understanding of what constitutes good regulation and the influence (if any) of different policy or geographic contexts. Moreover, it may prove beneficial to conduct some sort of ranking exercise to explore different stakeholder perceptions on the prioritisation of different criteria of good regulation. Indeed, health economists and policy makers may prioritise the pursuit of efficiency, industry may place value on effectiveness and efficiency, while citizens and politicians may emphasise the importance of accountability, transparency, and other procedural outcomes.

References

- Abelson, J., M. Giacomini, P. Lehoux and F. Gauvin (2007). Bringing “the public” into health technology assessment and coverage policy decisions: From principle to practice. *Health Policy* 82(1): 37–50.
- Abernethy, A.P., G. Raman, E.M. Balk, J.M. Hammond, L.A. Orlando, J.L. Wheeler, J. Lau and D.C. McCrory (2009). Systematic review: Reliability of compendia methods for off-label oncology indications. *Annals of Internal Medicine* 150: 336–343.
- Abraham, J. (1995). *Science, Politics and the Pharmaceutical Industry: Controversy and Bias in Drug Regulation*. London: University College London Press.
- Abraham, J. and G. Lewis (1998). Secrecy and transparency of medicines licensing in the EU. *Lancet* 352(9126): 480–482.
- Abraham, J. and G. Lewis (1999). Harmonising and competing for medicines regulation: How healthy are the European Union’s systems of drug approval? *Social Science and Medicine* 48(11): 1655–1667.
- Abraham, J. and G. Lewis (2000). *Regulating Medicines in Europe: Competition, Expertise and Public Health*. London: Routledge.
- Abraham, J. (2002). The pharmaceutical industry as political player. *Lancet* 360(9344): 1498–1502.
- Abraham, J. and G. Lewis (2002). Citizenship, medical expertise and the capitalist regulatory state in Europe. *Sociology* 36(1): 67–88.
- Abraham, J. and C. Davis (2005). A comparative analysis of drug safety withdrawals in the UK and US: Implications for current regulatory thinking and policy? *Social Science and Medicine* 61(5): 881–892.

Adamski, J. B. Godman, G. Ofierska-Sujkowska, et al. (2010). Risk-sharing arrangements for pharmaceuticals: Potential considerations and recommendations for European payer. *BMC Health Services Research* 10: 153.

AHRQ (1999). *The Outcome of Outcomes Research at AHCPR: Final report*. AHCPR Publication No. 99-RO44. Rockville, MD: AHRQ.

Alexander, G.C., L.P. Casalino and D.O. Meltzer (2003). Patient-physician communication about out-of-pocket costs. *Journal of the American Medical Association* 290(7): 953-958.

Altenstetter, C. (2003). EU and member state medical devices regulation. *International Journal of Technology Assessment in Health Care* 19(1): 228-248.

Altenstetter, C. and G. Permanand (2007). The EU regulation of medical devices and pharmaceuticals in comparative perspective. *Review of Policy Research* 24(5): 385-405.

Altenstetter, C. (2008). *Medical Devices: European Union Policymaking and the Implementation of Health and Patient Safety in France*. New Brunswick, NJ: Transaction Publishers.

Altenstetter, C. (2012). Medical device regulation in the European Union, Japan and the United States. Commonalities, differences and challenges. *Innovation: The European Journal of Social Science Research* 25(4): 362-388.

American Cancer Society (2010). *Cancer Facts & Figures 2010*. Atlanta: American Cancer Society.

Ancker, J.S. and D. Kaufman (2007). Rethinking health numeracy: A multidisciplinary literature review. *Journal of the American Medical Information Association* 14: 713-721.

- Anell, A. (2004). Priority setting in pharmaceuticals. *European Journal of Health Economics* 5: 28-35.
- Anell, A. and U. Persson (2005). Reimbursement and clinical guidance for pharmaceuticals in Sweden: Do health economic evaluations support decision making? *European Journal of Health Economics* 6(3): 274–279.
- Annemans, L., R. Crott, C.H. De, et al. (1997). Pricing and reimbursement of pharmaceuticals in Belgium. *Pharmacoeconomics* 11: 203-209.
- Appleby, J. (2008). The case of CT angiography: How Americans view and embrace new technology. *Health Affairs* 27(6): 1515–1521.
- Appleby, J., N. Devlin, D. Parkin, M. Buxton and K. Chalkidou (2009). Searching for cost-effectiveness thresholds in the NHS. *Health Policy* 91(3): 239–245.
- Atkinson, P. and A. Coffey (1997). Analysing documentary realities. *Qualitative Research: Theory, Method and Practice*. D. Silverman, ed. London: Sage Publications.
- Avorn, J. (2009). Debate about funding comparative effectiveness research. *New England Journal of Medicine* 360(19): 1927–1929.
- Avorn, J. and M. Fisher (2010). Bench to behavior: Translating comparative research into improved clinical practice. *Health Affairs* 29(10): 1891–1899.
- Audit Commission (2005). *Managing the Financial Implications of NICE Guidance*. London: Audit Commission.
- Australian Office of Regulatory Review (1998). *A Guide to Regulation*. Canberra: Commonwealth of Australia.
- Bach, P.B. (2009). Limits on Medicare’s ability to control rising spending on cancer drugs. *New England Journal of Medicine* 360(6): 626-633.

Bailey, K.D. (1994). *Methods of Social Research*. New York: The Free Press.

Balas, E. A. and S. A Boren (2000). Managing clinical knowledge for health care improvement. *Yearbook of Medical Informatics*. J. Bemmell, A.T. McCray, eds. Bethesda, MD: National Library of Medicine.

Baldwin, R., C. Scott, and C. Hood. (1998). Introduction. *A Reader on Regulation*. R. Baldwin, C. Scott, and C. Hood, eds. Oxford: Oxford University Press.

Baldwin, R. and J. Black (2007). *Really Responsive Regulation*. LSE Law, Society, and Economics Working Papers, 15-2007. London: London School of Economics and Political Science.

Baldwin, R., M. Cave, and M. Lodge (2010). *The Oxford Handbook of Regulation*. Oxford: Oxford University Press.

Baldwin, R. (2010). Better regulation: The search and the struggle. *The Oxford Handbook of Regulation*. R. Baldwin, M. Cave, and M. Lodge, eds. Oxford: Oxford University Press.

Baldwin, R., M. Cave and M. Lodge (2012). *Understanding Regulation: Theory, Strategy, and Practice*. Oxford: Oxford University Press.

Banta, D., W.J. Oortwijn, and W.T. Van Beekum (1995). *The Organisation of Health Care Technology Assessment in the Netherlands*. The Hague: Rathenau Institute.

Banta, D. and W. Oortwijn (2000). Health technology assessment and health care in the European Union. *International Journal of Technology Assessment in Health Care* 16(2): 626-635.

Banta, D. (2003). The development of health technology assessment. *Health Policy* 63: 121-132.

- Barbieri, D. and E. Ongaro (2008). EU agencies: What is common and what is distinctive compared with national-level agencies. *International Review of Administrative Sciences* 74(3): 395-420.
- Barbui, C., C. Baschiroto, and A. Cipriani (2011). EMA must improve the quality of its clinical trial reports. *British Medical Journal* 342: d2291.
- Barker, R. (2010). A flexible blueprint for the future of drug development. *Lancet* 374: 357-359.
- Barnato, A.E., M.B. Herdon, D.L. Anthony, P.M. Gallagher, J.S. Skinner, J.P.W. Bynum and E.S. Fisher (2007). Are regional variations in end-of-life care intensity explained by patient preferences? A study of the US Medicare population. *Medical Care* 45(5): 386–393.
- Barros, P. (2010). Pharmaceutical policies in European countries. *Advances in Health Economics and Health Services Research* 22: 3-27.
- Bassi, L., V. Bertele, and S. Garattini (2003). European regulatory policies in medicines and public health needs. *European Journal of Public Health* 13(3): 246-251.
- Basu, A. and T. J. Philipson (2010). *The Impact of Comparative Effectiveness Research on Health and Health Spending*. NBER Working Paper No. 15633. Cambridge, MA: National Bureau of Economic Research.
- Basu, S. and J.C. Hassenplug (2012). Patient access to medical devices – A comparison of U.S. and European review processes. *New England Journal of Medicine* 367: 485-488.
- Battista, R., J.M. Lance, P. Lehoux, and G. Regnier (1999). Health technology assessment and the regulation of medical devices and procedures in Quebec.

Synergy, collusion, or collision? *International Journal of Technology Assessment in Health Care* 15(3): 593-601.

Battista, R. and M.J. Hodge (1999). The evolving paradigm of health technology assessment: reflections for the Millennium. *Journal of the Canadian Medical Association* 160(10): 1464-1467.

Battista, R. (2006). Expanding the scientific basis of health technology assessment: A research agenda for the next decade. *International Journal of Technology Assessment in Health Care* 22(3): 275-282.

Becker, G.S., T.J. Philipson, and R.R. Soares (2007). *The Value of Life Near its End and Terminal Care*. National Bureau of Economic Research Working Paper 1333. Cambridge, MA: National Bureau of Economic Research.

Bending, M.W., J. Hutton and C. McGrath (2010). *Comparative-Effectiveness Versus Cost-Effectiveness: A Comparison of the French and Scottish Approaches to Single Technology Appraisal*. Paper presented at the ISPOR 15th Annual International Meeting, May 15–19th Atlanta, GA.

Berenson, A. (2005). Cancer drugs offer hope, but at a huge expense. *New York Times*, July 12.

Berg, M., T. van der Grinten, and N. Klazinga (2004). Technology assessment, priority setting and appropriate care in Dutch health care. *International Journal of Technology Assessment in Health Care* 20(1): 35-43.

Berndt, E. and J.P. Newhouse (2010). Pricing and reimbursement in U.S. pharmaceutical markets. *The Economics of the Biopharmaceutical Industry*. P.M. Danzon and S. Nicholson, eds. Oxford: Oxford University Press.

Bimber, B. (1996). *The Politics of Expertise in Congress: The Rise and Fall of the Office of Technology Assessment*. Albany: State University of New York Press.

- Black, J. and R. Baldwin R. (2010). Really responsive risk-based regulation. *Law and Policy* 32(2): 181-213.
- Blumenthal, D. (1983). Federal policy toward health care technology: The case of the National Center. *Milbank Quarterly* 61(4): 584–612.
- Blumstein, J.F. (1997). The Oregon experiment: The role of cost-benefit analysis in the allocation of Medicaid funds. *Social Science and Medicine* 45(4): 545–554.
- Bogner, A., B. Littig, and W. Menz (2009). *Interviewing Experts*. London: Palgrave Macmillan.
- Borowski, H.Z., J. Brehaut, and D. Hailey (2007). Linking evidence from health technology assessments to policy and decision making: The Alberta Model. *International Journal of Technology Assessment in Health Care* 23(2): 155-161.
- Bos, M. (2000). Health technology assessment in the Netherlands. *International Journal of Technology Assessment in Health Care* 16(2): 485-519.
- Boulenger, S., J. Nixon, M. Drummond, P. Ulmann, S. Rice, and G. De Pouvourville (2005). Can economic evaluations be made more transferable? *European Journal of Health Economics* 6: 334-346.
- Bovens, M. (2007). Analysing and assessing accountability: a conceptual framework. *European Law Journal* 13: 447. DOI: 10.1111/j.1468-0386.2007.00378.x
- Bowen, J.M., R. Hopkins, M. Chiu, G. Blackhouse, C. Lazzam, D. Ko, J.V. Tu, et al. (2007). *Clinical and Cost-Effectiveness Analysis of Drug Eluting Stents Compared to Bare Metal Stents for Percutaneous Coronary Interventions in Ontario: Final report*. HTA002-0705-02. Hamilton, Ontario, Program for Assessment of Technology in Health, St. Joseph's Healthcare Hamilton/McMaster University.
- Boyle, J.F. (1981). NCHCT could impede health care advancement and delivery. *Medical Instrumentation* 15(5): 296–298.

Braithwaite, J. (2006). Responsive regulation and developing economies. *World Development* 34(5): 884-898.

Breckenridge, A. (2010). Regulatory challenges, reimbursement, and risk-benefit assessment. *Clinical Pharmacology & Therapeutics* 88(2): 153-154.

Brickwood, D. (2004). *Implementation of NICE Guidance: An ABHI Perspective*. London: Department of Health.

Briggs, A., K. Ritchie, E. Fenwick, K. Chalkidou, P. Littlejohns. (2010). Access with evidence development in the UK. Past experience, current initiatives and future potential. *Pharmacoeconomics* 28: 163-170.

Brosheid, A. and D. Coen. (2007). Lobbying activity and fora creation in the EU: Empirically exploring the nature of the policy good. *Journal of European Public Policy* 14(3): 346-365.

Brown, L.D. (1991). The national politics of Oregon's rationing plan. *Health Affairs* 10(2): 28-51.

Brown, L.C., J.T. Powell, S.G. Thompson, D.M. Epstein, M. Sculpher, and R.M. Greenhalgh (2012). The UK EndoVascular Aneurysm Repair (EVAR) trials: Randomised trials of EVAR versus standard therapy. *Health Technology Assessment* 16(9): 1-218.

Bundesgesundheitsministerium (2008). *Stellungnahme zur Methodik der Kosten-Nutzen-Bewertung von Arzneimitteln*.
http://www.bmg.bund.de/cln_169/nn_1168258/SharedDocs/Standardartikel/DE/AZ/K/Glossar-Kosten-Nutzen-Bewertung/Stellungnahme.html. Bonn: Bundesgesundheitsministerium.

Burns, J. (2011). UnitedHealthcare's bold effort to deal with cancer drug costs. *Managed Care Magazine*, Jan 2011.

Busse, R., J. Orvain, M. Velasco, et al. (2002). Best practice in undertaking and reporting health technology assessments. *International Journal of Technology Assessment in Health Care* 18(2): 361-422.

Callahan, D. (2009). *Taming the beloved beast: How medical technology costs are destroying our health care system*. Princeton and Oxford: Princeton University Press.

Campillo-Artero, C. and F.M. Kovacs (2013). The use of risk-sharing tools for post adoption surveillance of non pharmacological technology in routine practice: results after one year. *BMC Health Services Research* 13: 181.

Canada External Advisory Committee on Smart Regulation (CEACSM) (2004). *Smart regulation: A Regulatory Strategy For Canada – Report To the Government of Canada*. Ottawa, External Advisory Committee on Smart Regulation.

Caniard, E. (2002). *Les Recommandations de Bonnes Pratique's: Un Outil de Dialogue, de Responsabilite' et de Diffusion de L'innovation*. Saint-Denis: Haute Autorite' de Sante'.

Cappellaro, G., S. Ghislandi, and E. Anessi-Pessina (2011). Diffusion of medical technology: The role of financing. *Health Policy* 100(1): 51-59.

Carlsson, P. (2004). Health technology assessment and priority setting for health policy in Sweden. *International Journal of Technology Assessment in Health Care* 14(4): 687-694.

Caron, M., M.P. Emery, P. Marquis, E. Piauult, and J. Scott (2008). Recent trends in the inclusion of patient-reported outcomes (PRO) data in approved drug labelling by the FDA and EMA. *Patient Reported Outcomes Newsletter* 40: 8-10.

Carpenter, D. and D.A. Moss (2013). *Preventing Regulatory Capture: Special Interest Influence and How to Limit It*. New York, NY: Cambridge University Press.

Carpenter, D., E.J. Zucker, and J. Avorn (2008). Drug-review deadlines and safety problems. *New England Journal of Medicine* 358: 1354-1361.

Carpenter, D. (2006). Reputation, gatekeeping and the politics of post-marketing drug regulation. *Virtual Mentor* 8: 403-406.

Carpenter, D. (2010). *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA*. Princeton, NJ: Princeton University Press.

Carman, K.L., M. Maurer, J. M. Yegian, J. McGee, M. Evers, and K.O. Marlow (2010). Evidence that consumers are sceptical about evidence-based health care. *Health Affairs* 29(7): 1400–1406.

Carino, T., R.D. Williams, A.M. Colbert, and P. Bridger (2006). Medicare's coverage of colorectal cancer drugs: A case study in evidence development and policy. *Health Affairs* 25(5): 1231-1239.

Carone, G., C. Schwierz, and A. Xavier (2012). *Cost-Containment Policies in Public Pharmaceutical Spending in the EU*. Economic Papers 461 (September). Brussels: European Commission.

Casson, C., F.J. Ruiz, and A. Miners (2013). How long has NICE taken to produce technology appraisal guidance? A retrospective study to estimate predictors of time to guidance. *BMJ Open*; 3:e001870 doi:10.1136/bmjopen-2012-001870.

Castigilione, D. (2006). Accountability. *Encyclopedia of Governance*. M. Bevir, ed. London: Sage.

Congressional Budget Office (CBO) (2007). *Research on the Comparative Effectiveness of Medical Treatments: Issues and Options for an Expanded Federal Role*. Washington, DC, Congressional Budget Office.

Cook, J., J. Vernon, and R. Mannin (2008). Pharmaceutical Risk-Sharing Agreements. *Pharmacoeconomics* 26: 551-556.

Cookson, R., D. McDaid, and A. Maynard (2001). Wrong SIGN, NICE mess: Is national guidance distorting allocation of resources. *British Medical Journal* 323: 743-745.

Cookson, R. and A. Maynard (2000). Health technology assessment in Europe: Improving clarity and performance. *International Journal of Technology Assessment in Health Care* 16(2): 639-650.

Council of the European Communities (CEC) (1965). Council Directive 65/65/EEC on the approximation of provisions laid down by law, regulation or administrative action relating to proprietary medicinal products. *Official Journal of the European Communities* 09 February 1965: L22/369.

Council of the European Communities (CEC) (1993). Council Regulation (EEC) No 2309/1993 of 22 July 1993 laying down community procedures for the authorization and supervision of medicinal products for human and veterinary use and establishing a European agency for the evaluation of medicinal products. *Official Journal of the European Communities* 24 August 1993: L214/01-21.

Council of the European Communities (CEC) (2001). Council Regulation 1049/01 regarding public access to European Parliament, Council and Commission documents. *Official Journal of the European Communities* 30 May 2001: L145 1.

Council of the European Communities (CEC) (2004). Regulation (EC) No 726/2004 of the European Parliament and of the Council of 31 March 2004 laying down community procedures for the authorization and supervision of medicinal products for human and veterinary use and establishing a European medicines agency. *Official Journal of the European Communities* 30 April 2004: L136/01-33.

Centers for Disease Control and Prevention (CDC) (2010). *National Vital Statistics Reports* 58(19). Atlanta: CDC, May 20.

Chalkidou, K., A. Hoy, P. Littlejohns (2007). Making a decision to wait for more research: When the National Institute for Health and Clinical Excellence recommends a technology only in the context of research. *J R Soc Med* 100(10): 453-460.

Chalkidou, K., J. Lord, A. Fisher and P. Littlejohns (2008). Evidence-based decision making: When should we wait for more information? *Health Affairs* 27: 1642–1623.

Chalkidou, K. (2009). *Comparative Effectiveness Review Within the UK's National Institute for Health and Clinical Excellence*. New York: The Commonwealth Fund.

Chalkidou, K., S. Tunis, R. Lopert, L. Rochaix, P. T. Sawicki, M. Nasser and B. Xerri (2009). Comparative effectiveness research and evidence-based health policy: Experience from four countries. *Milbank Quarterly* 87(2): 339–367.

Chalkidou, K. and T. Walley. (2010). Using comparative effectiveness research to inform policy and practice in the UK NHS: Past, present and future. *Pharmacoeconomics* 28(10): 799-811.

Chalkidou, K. (2012). Evidence and values: Paying for end-of-life drugs in the British NHS. *Health Economics, Policy and Law* 7(4): 393-409.

Chalkidou, K., S. Tunis, D. Wicher, R. Fowler, and M. Zwarenstein (2012). The role of pragmatic randomized controlled trials (pRCTs) in comparative effectiveness research. *Clinical Trials* 9(4): 436-446.

Chan A.W., A. Hrobjartsson, M.T. Haahr, P.C. Gotzsche, and D.G. Altman (2004). Empirical evidence for selective reporting of outcomes in randomized trials: Comparison of protocols to published articles. *Journal of the American Medical Association* 291: 2457-2465.

Chan, A.W. (2011). Access to clinical trial data. *British Medical Journal* 342:d80.

Chambers J.D., P.J. Neumann, and M.J. Buxton (2010). Does Medicare have an implicit cost-effectiveness threshold? *Medical Decision Making* 30(4): E14-E27.

Chen, C.E., S.S. Dhruva, L.A. Bero, and R.F. Redberg (2011). Inclusion of training patients in US Food and Drug Administration premarket approval cardiovascular device studies. *Arch Intern Med* 171: 534-539.

Cheng, M. (2003). *Medical Device Regulations: Global Overview and Guiding Principles*. Geneva, World Health Organization.

Chernew, M.E., A.B. Rosen, and A.M. Fendrick (2007). Value-based insurance design. *Health Affairs* 26(2): 195-203.

Chevreur, K. and I. Durand-Zaleski (2009). HTA in coverage and reimbursement decisions in France: Toward a new paradigm? *Euro Observer* 11(1): 5–6.

Chimowitz, M.I., M.J. Lynn, C. Derdeyn, et al. (2011). Stenting versus aggressive medical therapy for intracranial arterial stenosis. *New England Journal of Medicine* 365: 993-1003.

Chinitz, D. (2004). Health technology assessment in four countries: Response from political science. *International Journal of Technology Assessment in Health Care* 20(1): 55-60.

Chisholm, A., K. Briggs, and J. Askham (2009). *Not NICE: Can PCTs Engage Patients and the Public in Commissioning New Health Technologies?* Oxford: Picker Institute Europe.

Chiti, E. (2000). The emergence of a community administration: The case of European agencies. *Common Market Law Review* 37: 309-43.

Cholski, D.A., J. Avorn, and A.S. Kesselheim (2010). Designing comparative effectiveness research on prescription drugs: Lessons from the clinical trial literature. *Health Affairs* 29(10): 1842-1848.

Choudhry, N.K., M.B. Rosenthal, and A. Milstein (2010). Assessing the evidence for value-based insurance design. *Health Affairs* 29: 1988-1994.

Ciani, O. and R. Taylor (2013). Comparing drug and non-drug technologies in comparative effectiveness research. *Handbook of Health Services Research*. B. Sobolev, ed. New York: Springer.

Claxton, K., J.T. Cohen, and P. Neumann (2005). When is evidence sufficient? A framework for making use of all available information in medical decision making and for deciding whether more is needed. *Health Affairs* 24(1): 93-101.

Claxton, K. (2009). *Innovation Review – The Value of Innovation*. Presentation at Kennedy Workshop on NICE’s Approach to Valuing Innovation, London, May 19.

Claxton, K., S. Palmer, L. Longworth, et al. (2012). Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development. *Health Technology Assessment* 16(46): 1-323.

Clement, F., M.A. Harris, J.J. Li, K. Yong, K.M. Lee, and B.J. Manns (2009). Using effectiveness and cost-effectiveness to make drug coverage decisions: A comparison of Britain, Australia, and Canada. *Journal of the American Medical Association* 302: 1437-1443.

Clyde AT, L. Bockstedt, J.A. Farkas, and C. Jackson (2008). Experience with Medicare’s new technology add-on payment program. *Health Affairs* 27: 1632-1641.

CM/A (2000). *Evaluation of the Operation of Community Procedures for the Authorisation of Medicinal Products – Report Carried Out on Behalf of the European Commission by Cameron McKenna/Andersen Consulting*. European Commission: Directorate-General Enterprise (Pharmaceuticals and Cosmetics). Brussels: Commission of the European Communities.

CMS (Centers for Medicare and Medicaid Services) (1999). Procedures for making national coverage decisions. *Federal Register* 64(22), April 27.

CMS (Centers for Medicare and Medicaid Services) (2000). Criteria for making national coverage decisions. *Federal Register* 65(95), May 16.

CMS (Centers for Medicare and Medicaid Services) (2006). *Decision Memo for Intracranial Stenting and Angioplasty (CAG-0085R5)*. Baltimore, MD: Centers for Medicare and Medicaid Services).

CMS (Centers for Medicare and Medicaid Services) (2012a). *Medicare Program; Meeting of the Medicare Evidence Development and Coverage Advisory Committee*. Doc. No: 2012-6309, CMS-3261-N, 15372-15373. Baltimore, MD: Centers for Medicare and Medicaid Services).

CMS (Centers for Medicare and Medicaid Services) (2012b). *Decision Memo for Transcatheter Aortic Value Replacement (TAVR) (CAG-00430N)*. Baltimore, MD: Centers for Medicare and Medicaid Services).

CMTP (Center for Medical Technology Policy) (2011). *Coverage with Evidence Development: Key Issues*. Baltimore, MD: Center for Medical Technology Policy.

Code of Federal Regulations (2012). *CFR814.82*, April 1.

Cohen, D. and M. Billingsley (2011). European are left to their own devices. *British Medical Journal* 342: d2748.

Cohen, J. and W. Looney (2010). What is the value of oncology medicines? *Nature Biotechnology* 28: 1160-1163.

Coles, R. (1997). *Doing Documentary Work*. New York: Oxford.

College voor zorgverzekeringen (CVZ) (2012). *Conditional Reimbursement of Health Care*. Rapport Series No. 2012090234, April 6. Diemen, Netherlands: CVZ.

Committee Staff, House of Representatives Committee on Energy and Commerce (2011). *Internal memorandum*, Oct 14.

Conservatives for Patients' Rights (2009). *Conservatives for Patients' Rights Launches Multi-Million Dollar Campaign Promoting Free-Market Health Reform Solutions*. Press release, March 3, Conservatives for Patients' Rights.

Cookson, R., D. McDaid and A. Maynard (2001). Wrong SIGN, NICE mess: Is national guidance distorting allocation of resources. *British Medical Journal* 323: 743–745.

Cookson, R. and J. Hutton (2003). Regulating the economic evaluation of pharmaceuticals and devices: A European perspective. *Health Policy* 63(2): 167-178.

Cotter, D. (2009). The National Center for Health Technology: Lessons learned. *Health Affairs* (blog), January 22.

Council of Australian Governments. 2007. *Best Practice Regulation: A Guide for Ministerial Councils and National Standard Setting Bodies*. Oct. 2007.
http://www.finance.gov.au/obpr/docs/COAG_best_practice_guide_2007.pdf.

CRS (Congressional Research Service) (2007). *Comparative Clinical Effectiveness and Cost-Effectiveness: Background, History, and Overview*. Washington, DC: CRS.

CSDD (Center for the Study of Drug Development) (2004). *Faster Review, Longer Response Time Mark New EU Biopharmaceuticals*. Impact Reports – Analysis and Insight into Critical Drug Development Issues 6(2). Boston: Tufts University Center for the Study of Drug Development.

Cubbin, J. and J. Stern (2004). *Regulatory Effectiveness: The Impact of Good Regulation Governance on Electricity Industry Capacity and Efficiency in Developing Countries*. Working Papers 04/04, Department of Economics, City University. London: City University.

Culyer, A.J. (2006). NICE's use of cost effectiveness as an exemplar of a deliberative process. *Health Economics, Policy and Law* 1: 299-318.

Culyer, A.J. and J. Lomas (2006). Deliberative process and evidence-informed decision-making in health care: Do they work and how might be now. *Evidence and Policy* 2(3): 357-371.

Daemmrich, A.A. (2004). *Pharmacopolitics: Drug Regulation in the US and Germany*. Chapel Hill, NC: University of North Carolina Press.

Dahm, L.L. (2008). *The Elusive "Reasonable and Necessary": Making a Determination of Coverage with Evidence Development*.
[http://www.law.uh.edu/healthlaw/perspectives/\(LD\)CMSCoveragewithEvidence.pdf](http://www.law.uh.edu/healthlaw/perspectives/(LD)CMSCoveragewithEvidence.pdf).

Danzon, P.M. and E. Taylor (2010). Drug pricing and value in oncology. *Oncologist* 15(suppl 1): 24-31.

Darzi, A. (2008). *High Quality Care for All: NHS Next Stage Review Final Report*. London: UK Department of Health.

Department of Health (1999). *Faster Access to Modern Treatment: How NICE Appraisal Will Work*. London: Department Of Health.

Department of Health (2009). *Update of NICE Approved Cancer Drugs 2007/2008*. London: Department of Health.

Department of Health (2010). *Equity and Excellence: Liberating the NHS*. London: The Stationery Office Limited.

de Andres Trelles, F., S. Garattini, L. Gramstad, M. Johansson, F. Lekkerkerker, P. Rosi, C. Sampaio, B. Silva Lima, and M. Toivonen (2002). *Open Letter to Members of the European Parliament Regarding Proposed Amendments to the European Pharmaceutical Legislation*, Feb. 25.

http://www.haiweb.org/campaign/DTCA/MariaNegristatementDTCA_files/lettertoeu.htm.

DePanizza, A. (2007). *Composite and Decomposable Indicators for Evaluating RIA systems in Practice: Proposals for Discussion and Testing*. NBER Working Papers, Vol. 2007, No. 9: 1-21. Cambridge, MA: National Bureau of Economic Research.

Denscombe, M. (1998). *The Good Research Guide For Small Scale Social Research Projects*. Buckingham: Open University Press.

de Souza, J.A., M.J. Ratain, and A.M. Fendrick (2012). Value-based insurance design: Aligning incentives, benefits, and evidence in oncology. *Journal of the National Comprehensive Cancer Network* 10(1): 18-23.

Devlin, N., D. Parkin and M. Gold (2003). WHO evaluates NICE. *British Medical Journal*, 327: 1061–1062.

Devlin, N. and D. Parkin (2004). Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. *Health Economics* 13: 437–452.

Dhalla, I.A., S. Garner, K. Chalkidou, and P. Littlejohns (2009). Perspectives on the National Institute for Health and Clinical Excellence's recommendations to use health technologies only in research. *International Journal of Technology Assessment in Health Care* 25: 272-280.

Diel, P., W. Reuss, E. Aghayev, P. Moulin, C. Roder, and the SWISSspine Registry Group (2010). SWISSspine – a nationwide health technology assessment registry for balloon kyphoplasty: Methodology and first results. *Spine* 10(11): 961-971.

DiMaggio, P.J. and Powell W.W. (1991). Introduction. *The New Institutionalism in Organizational Analysis*. W.W. Powell and P.J. DiMaggio, eds. Chicago: University of Chicago Press.

Di Mario, C., S. James, D. Dudek, M. Sabate, and M. Degertekin (2011).

Commentary: The risk of over-regulation. *BMJ* 2011; 342: d3021.

DoBias, M. (2011). No death panels, please, but poll shows Americans can handle end-of-life chat. *National Journal*, March 8.

Dolgin, E. (2009). Publication bias continues despite clinical-trial registration. *Nature*, Sept. 3 doi:10.1038/news.2009.902.

Downing, N.S., J.A. Aminawung, N.D. Shah, J.B. Braunstein, H.M. Krumholz, and J.S. Ross (2012). Regulatory review of novel therapeutics – comparison of three regulatory agencies. *New England Journal of Medicine* 366: 2284-2493.

Draborg, E. and C.K. Andersen. (2005). What influences the choice of assessment methods in health technology assessments? Statistical analysis of international health technology assessments from 1989 to 2002. *International Journal of Technology Assessment in Health Care* 22(2): 155-160.

Draborg, E., D. Gyrd-Hansen, P.B. Poulsen, and M. Horder (2005). International comparisons of the definition and the practical application of health technology assessment. *International Journal of Technology Assessment in Health Care* 21: 89-95.

Dreyer, N.A., S.R. Tunis, M. Berger, D. Ollendorf, P. Mattox, and R. Gliklich (2010). Why observational studies should be among the tools used in comparative effectiveness research. *Health Affairs* 29(10): 1818-1825.

Drummond, M. and H. Weatherly. (2000). Implementing the findings of health technology assessments: If the CAT got out of the bag, can the TAIL wag the dog? *International Journal of Technology Assessment in Health Care* 6(1): 1-12.

Drummond, M. (2004). Economic evaluation in health care: Is it really useful or are we just kidding ourselves? *Australian Economic Review* 37(1): 3-11.

Drummond, M. and M.J. Sculpher (2005). Common methodological flaws in economic evaluations. *Medical Care* 43 (Supplement 7): 165-171.

Drummond, M., M.J. Sculpher, G.W. Torrence, B.J. O'Brien, and G.L. Stoddart (2005). *Methods for the Economic Evaluation of Health Care Programmes. Third Edition*. Oxford: Oxford University Press.

Drummond, M.F., J.S. Schwartz, B. Jonsson, B.R. Luce, P.J. Neumann, U. Siebert and S.D. Sullivan (2008). Key principles for the improved conduct of health technology assessments for resource allocation decisions. *International Journal of Technology Assessment in Health Care* 24: 244–258.

Drummond, M. (2009), More haste, less speed? The emerging practice of HTA in the United Kingdom. *Euro Observer*, 11(1): 9–10.

Drummond, M. and C. Sorenson (2009). Nasty or nice? A perspective on the use of health technology assessment in the United Kingdom. *Value in Health* 12 (Supplement S2: S8-S13).

Drummond, M.F., A.Griffin, and R. Tarricone R. (2009). Economic evaluation for devices and drugs. Same or different? *Value in Health* 12(4): 402-404.

Drummond, M. and A. Towse (2012). Is it time to consider the role of patient co-payments for pharmaceuticals in Europe? *European Journal of Health Economics* 13(1): 1-5.

Drummond, M. (2013). Twenty years of using economic evaluations for drug reimbursement decisions. What has been achieved? *Journal of Health Politics, Policy and Law* 38(6): 1081-1102.

Department of Health and Human Services (DHHS) (2009). *Adverse Event Reporting for Medical Devices*. Boston: Inspector General, US Department of Health and Human Services.

- Dhruva, S.S., L.A. Bero, and R.F. Redbury (2009). Strength of study evidence examined by the FDA in premarket approval of cardiovascular devices. *Journal of the American Medical Association* 302: 2679-2685.
- Drazen, J.M., J. Rainey, H. Begg, and A.S. Butler (2007). *Forum on Drug Discovery, Development, and Translation Adverse Drug Event Reporting: The Roles of Consumers and Health Care Professionals*. Washington, DC: The National Academies Press.
- Dunn, W.N. (2004). *Public Policy Analysis: An Introduction*. Third Edition. New Jersey: Prentice Hall.
- Earle, C.C., M.B. Landrum, J.M. Souza, B.A. Neville, J.C. Weeks, and J.Z. Ayanian (2008). Aggressiveness of cancer near the end of life: Is it a quality of care issue? *Journal of Clinical Oncology* 26: 3860-3866.
- Eberlein, B and D. Kerwer (2002). *Theorising the New Modes of European Union Governance*. European Integration Online Papers 6(5).
<http://eiop.or.at/eiop/texte/2002-005a.htm>.
- Eckermann, S. and A.R. Willan (2008). The option value of delay in health technology assessment. *Medical Decision Making* 28: 300-305.
- Eddy, D.M. (1991). Oregon's methods: Did cost-effectiveness analysis fail? *Journal of the American Medical Association* 266(15): 2135–2141.
- Egeberg, M. (2006). The institutional architecture of the EU and the transformation of European politics. *Multilevel Union Administration: The Transformation of Executive Politics in Europe*. M. Egeberg, ed. Basingstoke: Palgrave Macmillan.
- Eichler, H.G., B. Bloechl-Daum, E. Abadie, D. Barnett, F. Konig, and S. Pearson (2010). Relative efficacy of drugs: An emerging issue between regulatory agencies and third-party payers. *Nature Reviews* (April) 10: 277-291.

- Eicher, H.G., E. Abadie, A. Breckenridge, et al. (2011). Bridging the efficacy-effectiveness gap: A regulator's perspective on addressing variability of drug response. *Nature Reviews Drug Discovery* 10: 495-506.
- Eicher, H.G., K. Oye, L.G. Baird, et al. (2012). Adaptive licensing: Taking the next step in the evolution of drug approval. *Clin Pharmacol Ther* 91: 426-437.
- Eikermann, M., C. Gluud, M. Perleth, C. Wild, I. Gutierrez-Ibarluzea, S. Antoine, et al. (2013). Commentary: Europe needs a central, transparent, and evidence based regulation process for devices. *British Medical Journal* 346: f2771.
- Eisenberg, J. and D. Zarin (2002). Health technology assessment in the United States: Past, present, and future. *International Journal of Technology Assessment in Health Care* 18(2): 192–198.
- Elshaug, A., J.E. Hiller, S.R. Tunis, and T.R. Moss (2007). Challenges in Australian policy processes for disinvestment from existing, ineffective health care practices. *Australian New Zealand Health Policy* 4: 23.
- EMA (European Medicines Agency) (2000). *Sixth General Report on the Activities of the European Agency for the Evaluation of Medicinal Products*. London: European Medicines Agency.
- EMA (European Medicines Agency) (2001). *Seventh Annual Report*. London: European Medicines Agency.
- EMA (European Medicines Agency) (2002). *Eighth Annual Report*. London: European Medicines Agency.
- EMA (European Medicines Agency) (2003). *Ninth Annual Report*. London: European Medicines Agency.

EMA (European Medicines Agency) (2004). *Tenth Annual Report of the European Medicines Agency*, 2004. EMEA/61492/2005/EN/Final. London: European Medicines Agency.

EMA (European Medicines Agency) (2005). *Annual Report of the European Medicines Agency*, 2005. EMEA/MB/63019/2006/EN/Final. London: European Medicines Agency.

EMA ((European Medicines Agency) (2006a). *Rules for the Implementation of Regulation (EC) No. 1049/2001 on Access to EMEA Documents*.

EMEA/MB/203359/2006. London: European Medicines Agency.

www.ema.europa.eu/docs/en_GB/document_library/Other/2010/02/WC500070829.pdf.

EMA (European Medicines Agency) (2006b). *The European Medicines Agency Road Map to 2010: Preparing the Ground for the Future*. London: European Medicines Agency.

<http://www.EMA.eu.int/pdfs/general/direct/directory/3416303enF.pdf>.

EMA (European Medicines Agency) (2006c). *Annual Report of the European Medicines Agency*, 2006. EMEA/MB/24167/2007/EN/Final. London: European Medicines Agency.

EMA (European Medicines Agency) (2007). *Annual Report of the European Medicines Agency 2006*. EMEA/MB/24167/EN/FINAL London: European Medicines Agency.

EMA (European Medicines Agency) (2008). *Annual Report of the European Medicines Agency 2007*. EMEA/MB/17464/2008 London: European Medicines Agency.

EMA (European Medicines Agency) (2009). *Annual Report of the European Medicines Agency 2008*. EMEA/MB/330566/2009 London: European Medicines Agency.

EMA (European Medicines Agency) (2010a). European Medicines Agency Widens Public Access to Documents. Press release, 30 November.
www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2010/11news_detail_001158.jsp&mid=WC0b01ac058004d5c1&murl=menus/news_and_events/news_and_events.jsp.

EMA ((European Medicines Agency) (2010b). *Reflection Paper on the Need for Active Control in Therapeutic Areas Where Use of Placebo is Deemed Ethical and One or More Are Available*. EMA/759784/2010, November. London: European Medicines Agency.

EMA (European Medicines Agency) (2010c). *Annual Report 2008*. EMEA/MB/69923/2010 London: European Medicines Agency.

EMA (European Medicines Agency) (2011a). *Road Map to 2015: The European Medicines Agency's Contribution to Science, Medicines and Health*. London: European Medicines Agency.
www.ema.europa.eu/docs/en_GB/document_library/Report/2011/01/WC500101373.pdf

EMA (European Medicines Agency) (2011b). *Implementing the European Medicines Agency's Road Map to 2015: The Agency's Contribution to Science, Medicines, Health: "From Vision to Reality"*. London: European Medicines Agency, Oct. 6.
www.ema.europa.eu/docs/en_GB/document_library/Other/2011/10/WC500115960.pdf

EMA (European Medicines Agency) (2011c). *Annual Report 2010*. EMEA/MB/306870/2011 London: European Medicines Agency.

EMA (European Medicines Agency) (2012a). *Annual Report 2011: Overview of the Agency's Contribution to Science, Medicines and Health of the European Union*. London: European Medicines Agency.

EMA (European Medicines Agency) (2012b). *Guideline on Good Pharmacovigilance Practices (GVP)*. London: European Medicines Agency.

EMA (European Medicines Agency) (2013). *Annual Report 2012*. London: European Medicines Agency.

Epstein, S.S. (2005). *Cancer-Gate: How to Win the Losing War on Cancer*. Amityville, NY: Baywood.

Ernst and Young (2010). *Evaluation of the European Medicines Agency*. Brussels: European Commission.

Espin, J. and J. Rovira (2007). *Analysis of Differences and Commonalities in Pricing and Reimbursement Systems in Europe*. Brussels: Commission of the European Communities, DG Enterprise and Industry.

European Commission (Commission communication COM) (2001a). *Madelkein Report on Better Regulation*. Brussels: Commission of the European Communities.

European Commission (Commission communication COM) (2001b). *European Governance – A White Paper*. Brussels: Commission of the European Communities.

European Commission (Commission communication COM) (2002). *Simplifying and Improving the Regulatory Environment*. Brussels: Commission of the European Communities.

European Commission (Commission communication COM) (2003). *Updating and Simplifying the Community Acquis*. Brussels: Commission of the European Communities.

European Commission (Commission communication COM) (2006). *A Strategic Review of Better Regulation in the European Union*. Brussels: Commission of the European Communities.

European Commission (Commission communication COM) (2008). *Second Strategic Review of Better Regulation in the European Union*. Brussels: Commission of the European Communities.

European Commission (Commission communication COM) (2009). *Third Strategic Review of Better Regulation in the European Union*. Brussels: Commission of the European Communities.

European Commission (Commission communication COM) (2010a). *Smart Regulation in the European Union*. Brussels: Commission of the European Communities.

European Commission (Commission communication COM) (2010b). *Europe 2020: A European Strategy for Smart, Sustainable and Inclusive Growth*. Brussels: Commission of the European Communities.

European Commission (2011). *Exploring Innovative Healthcare: The Role of Medical Technology Innovation and Regulation*. Brussels: Commission of the European Communities.

European Commission (Commission communication COM) (2012a). *EU Regulatory Fitness*. Brussels: Commission of the European Communities.

European Commission (2012b). *Proposal for a Regulation of the European Parliament and of the Council on In Vitro Diagnostic Medical Devices*. Brussels: Commission of the European Communities. http://ec.europa.eu/health/medical-devices/files/revision_docs/proposal_2012_541_en.pdf.

European Parliament and Council (2001). *Regulation 1049/2001/EC Regarding Public Access to European Parliament, Council and Commission Documents, OJ 2001 No. L145/43*. Brussels: European Parliament and Council.

European Parliament and Council (2011). *Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the Application of Patients' Rights in Cross-Border Healthcare*. Brussels: European Parliament and Council.

External Advisory Committee on Smart Regulation (2004). *Smart Regulation: A Regulatory Strategy for Canada*. Canada: External Advisory Committee on Smart Regulation. www.smartregulation.gc.ca.

Facey, K., A. Boivin, J. Gracia, H.P. Hansen, A. Lo Scalzo, J. Mossman, and A. Single (2010). Patients' perspectives in health technology assessment: A route to robust evidence and fair deliberation. *International Journal of Technology Assessment in Health Care* 26(3): 334-340.

Facey, K. (2011). Patient involvement in HTA: What added value? *Pharmaceuticals Policy and Law* 13(3-4): 245-251.

Faden, R.R., K. Chalkidou, J. Appleby, H.R. Waters, and J. Leider (2009). Expensive cancer drugs: A comparison between the United States and the United Kingdom. *Milbank Quarterly* 87: 789-819.

Faguet, G.B. (2005). *The War on Cancer: An Anatomy of Failure, A Blueprint for the Future*. Dordrecht, the Netherlands: Springer.

Federal Register (2012). *Food and Drug Administration Safety and Innovation Act (S.3187)*. Government Printing Office: Washington, DC.
<http://www.gpo.gov/fdsys/pkg/PLAW-112publ144/pdf/PLAW-112publ144.pdf>

Fenwick, E., K. Claxton, and M. Schulpher (2008). The value of implementation and the value of information: Combined and uneven development. *Medical Decision Making* 28(1): 21-32.

Fink, C., A. Mattoo, and R. Rathindran (2003). An assessment of telecommunications reform in developing countries. *Information Economics and Policy* 15(4): 443-66.

Flick, U. (2014). *An Introduction to Qualitative Research*. Fifth Edition. London: Sage.

Frederick, J. (2008). One sure election outcome: Change in healthcare system. *Drug Store News*, August 18.

Fojo, T. and C. Grady (2009). How much is life worth: Cetuximab, non-small cell lung cancer and the \$440 billion question. *Journal of the National Cancer Institute* 101: 1044-1048.

Folino-Gallo, P., S. Montilla, M. Bruzzone, and N. Martini (2008). Pricing and reimbursement of pharmaceuticals in Italy. *European Journal of Health Economics* 9(3): 305-310.

Foote, S. B. (2002). Why Medicare cannot promulgate a national coverage rule: A case of regula mortis. *Journal of Health Politics, Policy and Law* 2(5): 707–730.

Foote, S.B., D. Wholey, T. Rockwood, and R. Halpern (2004). Resolving the tug-of-war between Medicare's national and local coverage. *Health Affairs* 23(4): 108-123.

Forrest, S., M.M. Goetghebeur, and J. Hay (2005). Medicare coverage and reimbursement of outpatient prescription drugs in the US: History, recent changes and outlook for the future. *Applied Health Economics Health Policy* 4(1): 9-14.

Franken, M., M. le Polain, I. Cleemput, and M. Koopmanschap (2012). Similarities and differences between five European drug reimbursement systems. *International Journal of Technology Assessment in Health Care* 28(4): 349-357.

Fraser, A.G., M.W. Krucoff, R. Brindis, M. Komajda, and S.C. Smith (2011). International collaboration needed on device clinical standards. *British Medical Journal* 342: d2952.

Freemantle, N. (2011). Evaluating and regulating the device industry. *British Medical Journal* 342: d2839.

Freemantle, N. and M.F. Drummond (2010). AHRQ series commentary 3: The United States addresses comparative effectiveness but not cost-effectiveness through the effective health-care program. *Journal of Clinical Epidemiology* 63: 479-480.

Fronsdal, K., F. Pichler, L. Mardhani-Bayne, C. Henshall, J.A. Rottingen, B. Morland, and M. Klemp (2012). Interaction initiatives between regulatory, health technology assessment and coverage bodies, and industry. *International Journal of Technology Assessment in Health Care* 28(4): 374-381.

Furberg, C.D., A.A. Levin, P.A. Gross, R.S. Shapiro, and B.L. Storm (2006). The FDA and drug safety: A proposal for sweeping changes. *Archives of Internal Medicine* 166(18): 1938-1942.

GAO (General Accounting Office) (1995). *Practice Guidelines: Overview of Agency for Health Care Policy and Research Efforts*. Washington, DC: General Accounting Office.

Garau, M., K.K. Shah, A.R. Mason, Q. Wang, A. Towse, and M.F. Drummond (2010). *Using QALYs in Cancer: A Review of the Methodological Limitations*. Office of Health Economics, Research Paper 10/01, Oct 2010. London: Office of Health Economics.

Garber, A.M. and H.C. Sox (2010). The role of costs in comparative effectiveness research. *Health Affairs* 29(10): 1805-1811.

Garattini, S. and V. Bertele (2004). The role of the EMA in regulating pharmaceutical products. *Regulating Pharmaceuticals in Europe: Striving for Efficiency, Equity and Quality*. E. Mossialos, M. Mrazek, and T. Walley, eds. Maidenhead: Open University Press.

Garattini, S. (2005). EMA: For patients or for industry? *Pharmacoeconomics* 23(3): 207-208.

Garattini, S. (2007). How can we regulate medicines better? *BMJ* 335(7624): 803-805.

Garattini, S. and I. Chalmers (2009). Patients and the public deserve big changes in evaluation of drugs. *British Medical Journal* 338: b1025.

Garattini, S. and V. Bertele (2010). Europe's opportunity to open up drug regulation. *British Medical Journal* 340: 842-843.

Garcia-Atlas, A., S. Onategui-Parra, and P. Neumann. (2004). Cross-national comparison of technology assessment processes. *International Journal of Technology Assessment in Health Care* 20(3): 200-310.

Gardner, J.S. (1996). The European agency for the evaluation of medicines and European regulation of pharmaceuticals. *European Law Journal* 2(1): 48-82.

Gatyas, G. and L. Longwell (2008). IMS Health reports U.S. prescription sales grew 3.8 percent in 2007, to \$286.5 billion. Press release, March 12. Norwalk, CT: IMS Health.

Garrison, L., A. Towse, A. Briggs, G. de Pouvourville, J. Grueger, P. Mohr, J.L. Severens, P. Siviero, and M. Sleeper (2013). Performance-based risk-sharing arrangements - Good practices for design, implementation and evaluation: ISPOR Good Practices for Performance-Based Risk-Sharing Arrangements Task Force Report. *Value in Health* 16(5): 703-719.

Gauvin, F.P., J. Abelson, M. Giacomini, J. Eyles, and J.N. Lavis (2010). "It all depends": Conceptualizing public involvement in the context of health technology assessment agencies. *Social Science and Medicine* 70(10): 1518-1526.

Gawande, A. (2010). Letting go: What should Medicine do when it can't save your life? *The New Yorker*, Aug. 2.

Gehring, T. and S. Krapohl (2007). Supranational regulatory agencies between independence and control: The EMEA and the authorization of pharmaceuticals in the European single market. *Journal of European Public Policy* 14(2): 208-226.

Gelijns, A.C., L.D. Brown, C. Magnell, E. Ronchi, and A.J. Moskowitz (2005). Evidence, politics, and technological change. *Health Affairs* 24(1): 29-40.

Gelijns, A.C., M.J. Russo, K.N. Hong, L.D. Brown, D.D. Ascheim, and A.J. Moskowitz (2013). Dynamics of device innovation: Implications for assessing value. *International Journal of Technology Assessment in Health Care* 29(4): 365-373.

Genoud, C. (2003). *Regulation as a Game: The Role of Independent Regulatory Agencies in the Regulatory Process*. London: London School of Economics.
<http://www.lse.ac.uk/researchAndExpertise/units/CARR/pdf/ChristopheGenoud.pdf>.

Geradin, D., R. Muñoz, and N. Petit (eds) (2005). *Regulation through Agencies in the EU: A New Paradigm of European Governance*. Cheltenham: Edward Elgar Publishing Limited.

Gerber, A.S., E.M. Patashnik, D. Doherty, and C. Cowling (2010). The public wants information, not broad mandates, from comparative effectiveness research. *Health Affairs* 29(10): 1872-1881.

Gerber, A.S. and E.M. Patashnik (2010). Problem solving in a polarized age: Comparative effectiveness and the politicization of evidence-based medicine. *Forum Manuscript* 353: 1-13.

Geradin, D. and N. Petit (2004). *The Development of Agencies at EU and National Levels: Conceptual Analysis and Proposals for Reform*. Jean Monnet Working Paper 01/04. New York School of Law.

Gerhardus, A. (2006). The role of HTA in German health care. Do we need impact objectives? *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 49(3): 233–240.

Giezen, T.J., A.K. Mantel-Teeuwisse, S. Straus, H. Schellekens, H. Leufkens and A.C.G. Egberts (2008). Safety-related regulatory actions for biologicals approved in the United States and the European Union. *Journal of the American Medical Association* 300(16): 1887–1896.

Gilardi F. (2002). Policy credibility and delegation to independent regulatory agencies: A comparative empirical analysis. *Journal of European Public Policy* 9(6): 873-893.

Gilardi F. (2003). *Evaluating Independent Regulators. OECD Working Party of Regulatory Management and Reform Document 15*. Paris: Organization for Economic Co-operation and Development.

Gilardi F. (2005). The institutional foundations of regulatory capitalism: The diffusion of independent regulatory agencies in Western Europe. *Annals of the American Academy of Political and Social Science* 598: 84-101.

Gilardi, F., R. Munoz, and N. Petit (eds). (2005). *Regulation Through Agencies in the EU: A New Paradigm of European Governance*. Cheltenham: Edward Elgar Publishing Limited.

Gillick, M.R. (2004). Medicare coverage for technological innovations: Time for new criteria? *New England Journal of Medicine* 350(21): 2199-2203.

Goddard, M., K. Hauk, A. Preker and P. Smith (2006). Priority-setting in health: A political economy perspective. *Health Economics, Policy and Law* 1: 79–90.

Goeree, R. and L. Levin (2006). Building bridges between academic research and policy formulation. *Pharmacoeconomics* 24(11): 1143–1156.

Goldman, D.P., A.B. Jena, D.N. Lakdawalla, J.L. Malin, J.D. Malkin and E. Sun (2010). The value of specialty oncology drugs. *Health Services Research* 45(1): 115–132.

Goodman, C. (2004). *HTA 101: Introduction to Health Technology Assessment*. Bethesda, MA: National Library of Medicine, National Information Center on Health Services Research and Health Care Technology.

Goodman, D.C., E.S. Fisher, C. Chang, N.E. Morden, J.O. Jacobson, K. Murray and S. Miesfeldt (2010). *Quality of End-of-Life Cancer Care for Medicare Beneficiaries: Regional and Hospital-Specific Analyses: A Report of the Dartmouth Atlas Project*. Lebanon, NH: Dartmouth Institute.

Goodman, D.C., E.S. Fisher, C. Chang, N.E. Morden, J.O. Jacobson, K. Murray and S. Miesfeldt (2011). *Quality of End-of-Life Cancer Care for Medicare Beneficiaries: Regional and Hospital-Specific Analyses: A Report of the Dartmouth Atlas Project*. Lebanon, NH: Dartmouth Institute.

Goodman, D. (2011). End-of-life cancer care in Ontario and the United States: Quality by accident or quality by design? *Journal of the National Cancer Institute* 103(11): 1–2.

Gottlieb, S. (2011). How the FDA could cost you your life. *Wall Street Journal*, October 3.

Gotzsche, P.C. and A.W. Jorgensen (2011). Opening up the European Medicines Agency. *BMJ* 342: d2686.

Gray, B.H., M.K. Gusmano, and S. Collins (2003). AHCPR and the politics of health services research. *Health Affairs* (Web Exclusive), June 25.

Greenberg, D., W.C. Winkelmayr, and P. Neumann (2005). Prevailing judgements about society's willingness to pay for QALY or life-year gained. *International Journal of Technology Assessment in Health Care* 2 (Suppl 1): 301.

Groenleer, M. (2009). *The Autonomy of European Union Agencies: A Comparative Study of Institutional Development*. Leiden, Netherlands: Proefschrift, University of Leiden.

Guillou, P.J., P. Quirke, H. Thorpe, et al. (2005). Short-term endpoints of conventional versus laparoscopic-assisted surgery in patients with colorectal cancer (MRC CLASICC trial): Multicentre, randomised controlled trial. *Lancet* 365: 1718–1726.

Gulacsi, L., T. David, and C. Dozsa (2002). Pricing and reimbursement of drugs and medical devices in Hungary. *European Journal of Health Economics* 3(4): 271-278.

Gusmano, M.K. and B.H. Gray (2010). *Evidence and Fear: Navigating the Politics of Evidence Based Medicine*. AcademyHealth Reports, June. Washington, DC: AcademyHealth.
<http://www.academyhealth.org/files/publications/academyhealthreports/evidenceandfear.pdf>

Haga, A. and J.M. Sverre (2002). Pricing and reimbursement of pharmaceuticals in Norway. *European Journal of Health Economics* 3(3): 215-220.

Hanrelty, C. and C. Koop (2012). Measuring the formal independence of regulatory agencies. *Journal of European Public Policy* 19(2): 198-216.

Harrington, S.E. and T.J. Smith (2008). The role of chemotherapy at the end of life: When is enough, enough? *Journal of the American Medical Association* 299: 2667–2678.

Harris, A.H., M. Buxton, B. O'Brien, F. Rutten, and M. Drummond (2001). Using economic evidence in reimbursement decisions for health technologies: Experience of 4 countries. *Expert Rev Pharmacoeconomics Outcome Res* 1: 7-12.

Haycox, A. (2008). Does NICE blight exist, and if so, why?" *Pharmacoeconomics* 26(12): 987–989.

Health Action International (HAI) (2005). *Does the European Patients Forum Represent Patients' or Industry Interests? A Case-Study in the Need for Financial Disclosure*. Amsterdam: Health Action International.

Health Care Financing Administration (HCFA) (1989). Medicare program: Criteria and procedures for making medical services coverage decisions that relate to health care technology. *Federal Register* 54(30), January 30.

Healy, P. and C. Dean. (2007). Clinton says she would shield science from politics. *New York Times*, October 5.

Hecht, J.R., E. Mitchell, T. Chidiac, et al. (2009). A randomized phase IIIB trial of chemotherapy, bevacizumab, and panitumumab compared with chemotherapy and bevacizumab alone for metastatic colorectal cancer. *Journal of Clinical Oncology* 27(5): 672–680.

Heneghan, C. (2011). The saga of poly implant prostheses breast implants. *BMJ* 344: e306.

Henshall, C., L. Sansom, H. Eichler, A. Lemgruber, C. Longson, B. O'Rourke, and S. Tunis (2013). Understanding the role and evidence expectations of health technology assessment and coverage/payer bodies. What are they looking for, and how and why does this differ from what regulators require? *Therapeutic Innovation & Regulatory Science* doi:10.1177/21684479013512488.

Henshall, C., L. Mardani-Bayne, K.B. Fronsdal, and M. Klemp (2011). Interactions between health technology assessment, coverage, and regulatory processes: Emerging issues, goal, and opportunities. *International Journal of Technology Assessment in Health Care* 27(3): 253-260.

Herxheimer A. (2003). Relationships between the pharmaceutical industry and patients' organisations. *BMJ* 326: 1208-1210.

Himmelstein, D.U., D. Thorne, E. Warren and S. Woolhandler (2009). Medical bankruptcy in the United States, 2007: Results of a national study. *American Journal of Medicine* 122: 741–746.

Hines, J.Z., P. Lurie, E. Yu, and S. Wolfe (2010). Left to their own devices: Breakdowns in United States medical device premarket review. *PLoS Med* 7(7): e1000280.

Hitchen, L. (2008). NICE recommendations have had little effect on multiple sclerosis services five years on. *British Medical Journal* 337: a734.

Hivon, M., P. Lehoux, J.L. Denis, and S. Tailliez (2005). Use of health technology assessment in decision making: Coresponsibility of users and producers? *International Journal of Technology Assessment in Health Care* 21(2): 268-275.

Hollingsworth, J.M., Y. Zhang, S.L. Krein, Z. Ye and B.K. Hollenbeck (2010). Understanding the variation in treatment intensity among patients with early stage bladder cancer. *Cancer* 378: 3587–3594.

Hood, C. (2000). Paradoxes of public-sector managerialism, old public management and public service bargains. *International Public Management Journal* 3(1): 1-22.

Hood, C., H. Rothstein, and R. Baldwin (2001). *The Government of Risk: Understanding Risk Regulation Regimes*. Oxford: Oxford University Press.

Hsu, J., M. Price, J. Huang, R. Brand and V. Fung (2006). Unintended consequences of caps on Medicare drug benefits. *New England Journal of Medicine* 354: 2349–2359.

Hughes, B. (2008). Payers' growing influence on R&D decision making. *Nature Reviews Drug Discovery* 7: 876–878.

Huskamp, H.A., N.L. Keating, J.L. Malin, et al. (2009). Discussions with physicians about hospice among patients with Metastatic lung cancer. *Archives of Internal Medicine* 169: 954–962.

Hutton, J., C. McGrath, J. Frybourg, M. Tremblay, E. Bramley-Harker and C. Henshall (2006). Framework for describing and classifying decision-making systems using technology assessment to determine the reimbursement of health technologies (fourth hurdle systems). *International Journal of Technology Assessment in Health Care* 22: 10–18.

Hutton, J., P. Trueman, C. Henshall (2007). Coverage with evidence development: An examination of conceptual and policy issues. *International Journal of Technology Assessment in Health Care* 23: 425-435.

Hutton, J., P. Trueman, and K. Facey (2008). Harmonization of evidence requirements for health technology assessment in reimbursement decision making. *International Journal of Technology Assessment in Health Care* 24(4): 511-517.

Iglehart, J.K. (2010). The political fight over comparative effectiveness research. *Health Affairs* 29(10): 1757–1760.

Institute of Medicine (IOM) (2007). *Learning What Works Best: The Nation's Need for Evidence on Comparative Effectiveness in Health Care, IOM Roundtable on Evidence-Based Medicine*. Washington, DC: Institute of Medicine.

Institute of Medicine (IOM) (2009). *Initial National Priorities for Comparative Effectiveness Research, Report Brief, June 2009*. Washington, DC: Institute of Medicine, http://books.nap.edu/openbook.php?record_id512648.

Institute of Medicine (IOM) (2011). *Medical Devices and the Public's Health. The FDA 510(k) Clearance Process at 35 Years*. Washington, DC: National Academies Press. www.iom.edu/Reports/2011/Medical-Devices-and-the-Publics-Health-the-FDA-510k-Clearance-Process-at-35-Years.aspx.

Institute for Quality and Efficiency in Healthcare (IQWiG) (2008). *General Methods*. Cologne, Germany: IQWiG.

http://www.iqwig.de/download/IQWiG_General_methods_V-3-0.pdf.

Institute for Quality and Efficiency in Healthcare (IQWiG) (2010a). *About*. Cologne, Germany: IQWiG. <http://www.iqwig.de/institute-for-quality-and-efficiency-in-health.2.en.html>.

Institute for Quality and Efficiency in Healthcare (IQWiG) (2010b). *IQWiG in Dialogue*. Cologne, Germany: IQWiG. <http://www.iqwig.de/iqwig-in-dialogue.908.en.html>.

International Society of Drug Bulletins (2006). *About the Proposals of the EU Commission Regarding Drug Approvals and Pharmacovigilance – An opinion of the International Society of Drug Bulletins*. Paris: International Society of Drug Bulletins. http://66.71.191.169/isdbweb/pag/eu_comm.php.

Jacob, R. and M. McGregor (1997). Assessing the impact of health technology assessment. *International Journal of Technology Assessment in Health Care* 13(1): 68–80.

Jacobson, M., A.J. O'Malley, C.C. Earle, J. Pakes, P. Gaccione and J.P. Newhouse (2006). Does reimbursement influence chemotherapy treatment for cancer patients? *Health Affairs* 25: 437–443.

Jaramillo, A. (2013). *Benchmarking Governance as a Tool for Promoting Change*. World Bank Policy Research Working Paper No. 81050. Washington, DC: The World Bank Group.

Jeffreys, B. (2007). Fears over NHS cancer drug costs. *British Broadcasting Company*, May 14.

James, S., J.C. Daubert, and F. Van de Werf (2011). Use of registries to investigate the past and develop the future. *British Medical Journal* 342: d2826.

Johannsen, K.S., L.H. Pedersen, and E.M. Sorensen (2004). *Independent Regulatory Authorities – A Comparative Study of European Energy Regulators*. Copenhagen: Institute of Local Government Studies, University of Copenhagen.

Johnson, F.R. (2009). Moving the QALY forward or just stuck in traffic? *Value in Health* 12: S38–S39.

Jonsson, E. (2002). Development of health technology assessment in Europe. *International Journal of Technology Assessment in Health Care* 18: 171-183.

Juurlink, D.N. (2010). “Rosiglitazone and the case for safety over certainty. *Journal of the American Medical Association* 304(4): 469-471.

Kanavos, P. (2003). Overview of pharmaceutical pricing and reimbursement regulation in Europe. *Japanese Pharmacology and Therapeutics* 31(10): 819-838.

Kanavos, P., E. Nicod, S. Van den Aardweg and S. Pomedli (2010) “The Impact of Health Technology Assessments: An International Comparison” *Euro Observer* 12(4): 1–6.

Kane, N.M. (1997). Pharmaceutical cost containment and innovation in the United States. *Health Policy* 41(Supplement 1): S71-S89.

Kaiser Family Foundation (2012). *Health Care Costs: A Primer, Key Information on Health Care Costs and Their Impact*. Menlo Park, California: Kaiser Family Foundation.

Kaiser Family Foundation (2012). *Employee Health Benefits 2012 Annual Survey*. Menlo Park, California: Kaiser Family Foundation.

Kaufmann, D., A. Kraay, and M. Mastruzzi (1999). *Governance Matters*. World Bank Policy Research Working Paper No. 2195. Washington, DC: The World Bank Group.

Kaufmann, D., A. Kraay, and M. Mastruzzi (2002). *Governance Matters II: Updated Indictors for 2000/1*. World Bank Policy Research Working Paper No. 2772. Washington, DC: The World Bank Group.

Kaufmann, D., A. Kraay, and M. Mastruzzi (2004). *Governance Matters III: Governance Indicators for 1996-2002*. *World Bank Economic Review* 18: 253-287.

Kaufmann, D., A. Kraay, and M. Mastruzzi (2005). *Governance Matters IV: Governance Indicators for 1996-2004*. World Bank Policy Research Working Paper No. 3630. Washington, DC: The World Bank Group.

Kaufmann, D., A. Kraay, and M. Mastruzzi (2009). *Governance Matters VIII: Aggregate and Individual Governance Indicators 1996-2008*. World Bank Policy Research Working Paper No. 4978. Washington, DC: The World Bank Group.

Keating, N.L., M.B. Landrum, S.O. Rogers, et al. (2010). Physicians factors associated with discussions about end-of-life care. *Cancer* 116: 998–1006.

Keenan, P.S., P.J. Neumann, and K.A. Phillips (2006). Biotechnology and Medicare's new technology policy: Lessons from three case studies. *Health Affairs* 25(5): 260–269.

Kennedy, I. (2009). *Appraising the Value of Innovation and Other Benefits: a Short Study for NICE*. London: National Institute for Health and Clinical Excellence.

Kim, P. (2007). Costs of cancer: The patient perspective. *Journal of Clinical Oncology* 25(2): 228–232.

Kirisits, A. and W.K. Redekop (2013). The economic evaluation of medical devices. *Applied Health Economics and Health Policy* 11(1): 15-26.

Klein, R. (1997). Learning from others: Shall the last be the first? *Journal of Health Politics, Policy and Law* 22(5): 1267-1278.

Klemp, M., K.B. Fronsdal, and K. Facey (2011). What principles should govern the use of managed entry agreements? *International Journal of Technology Assessment in Health Care* 27(1): 77-83.

Knill, C. and A. Lenschow (2003). Modes of regulation in the governance of the European Union: Towards a comprehensive evaluation. *European Integration Online Papers* 7(1): 1-20.

Kolata, G. and A. Pollack (2008). The evidence gap: Costly cancer drug offers hope, but also a dilemma. *New York Times*, July 6.

Kramer, D.B., S. Xu, and A.S. Kesselheim (2012). Regulation of medical devices in the United States and Europe Union. *New England Journal of Medicine* 366(9): 848-855.

Kramer, D.B., Y.T. Tan, C. Sato, and A.S. Kesselheim (2013). Postmarket surveillance of medical devices: A comparison of strategies in the US, EU, Japan, and China. *PLOS Medicine* doi: 10.1271/journal.pmed.1001519.

Kraphol, S. (2004). Credible commitment in non-independent regulatory agencies: A comparative analysis of the European agencies for pharmaceuticals and foodstuffs. *European Law Journal* 10(5): 518-538.

Kristensen, F.B., M. Makela, S.A. Neikter, et al. (2009a). European network for health technology assessment, EUnetHTA: planning, development, and implementation of a sustainable European network for health technology assessment. *International Journal for Technology Assessment in Health Care* 25(Supplement 2): 107-116.

Kristensen, F.B., K. Lampe, D.L. Chase, et al. (2009b). Practical tools and methods for health technology assessment in Europe: Structures, methodologies, and tools developed by the European Network for Health Technology Assessment, EUnetHTA. *International Journal for Technology Assessment in Health Care* 25(Supplement 2): 1-8.

Kruger, K. And M. Kruger (2012). The medical device sector. *The Business of Healthcare Innovation*. L.R. Burns, ed. Cambridge, UK: Cambridge University Press.

Kunkle, G.C. (1995). New challenge or the past revisited? The Office of Technology Assessment in historical context. *Technology in Society* 17(2): 175–196.

Kupersmith, J. (2009). *What Can the Health Care System Learn from 30 Years of Comparative Effectiveness Research in the VA?* Paper presented at the Department of Veterans Affairs Symposium, October 22, Washington, DC.

Ladin, Z. and M. Imhoff (2010). *FDA Review Patterns of De Novo Submissions*. Boston: Boston MedTech Advisors.

Lakdawalla, D.N., D.P. Goldman, P.C. Michaud, et al. (2009). U.S. pharmaceutical policy in a global marketplace. *Health Affairs* 28(1): w138-w150.

Langa, K.M., A.M. Fedrick, M.E. Chernew, M.U. Kabeto, K.L. Paisley and J.A. Haymen (2004). Out-of-pocket health care expenditures among older Americans with cancer. *Value Health* 7(2): 186–194.

Latif, A. (2013). Value-based pricing. *Lancet UK Policy Matters*.
ukpolicymatters.thelancet.com/value-based-pricing/

Lauslahti, K., R. Roine, V. Semberg, M. Kekomaki, M. Kontinen, and P. Karp (2000). Health technology assessment in Finland. *International Journal of Technology Assessment in Health Care* 16(2): 382-399.

Lexchin, J. and O. O'Donovan (2010). Prohibiting or Managing conflict of interest? A review of policies and procedures in three European drug regulation agencies. *Social Science and Medicine* 70(5): 643-647.

Lee, N. and C. Kirkpatrick (2006). Evidence-based policy making in Europe: An evaluation of European commission integrated impact assessments. *Impact Assessment and Project Appraisal* 24(1): 23-33.

Lehoux, P. and S. Blume (2000). Technology assessment and the sociopolitics of health technologies. *Journal of Health Politics, Policy, and Law* 25(6): 1083-1120.

Lenzer, J. and S. Brownlee (2010). Why the FDA can't protect the public. *BMJ* 341: c4753.

Levin, L., R. Goeree, N. Sikich, B. Jorgensen, M.C. Brouwers, T. Easty, and C. Zahn (2007). Establishing a comprehensive continuum from an evidentiary base to policy development for health technologies: The Ontario experience. *International Journal of Technology Assessment in Health Care* 23(3): 299-309.

Levin, L., R. Goeree, M. Levine, M. Krahn, T. Easty, and D. Henry (2011). Coverage with evidence development: The Ontario experience. *International Journal of Technology Assessment in Health Care* 27: 159-168.

Lewis, G. and J. Abraham (2001). The creation of neo-liberal corporate bias in transnational medicines control: The industrial shaping and interest dynamics of the European regulatory state. *European Journal of Political Research* 39: 53-80.

Lexchin, J. and B. Mintzes (2008). Medicines reimbursement recommendations in Canada, Australia, and Scotland. *American Journal of Managed Care* 14: 581-588.

Lilford, R.J., D.A. Braunholtz, R. Greenhalgh, and S.J.L. Edwards (2000). Trials and fast changing technologies: the case for tracker studies. *BMJ* 320(7226): 43-46.

Lodge, M. (2004). Accountability and transparency in regulation: Critiques, doctrines and instruments. *The Politics of Regulation: Institutions and Regulatory Reforms for the Age of Governance*. J Jordana, D Levi-Faur, eds. Cheltenham: Edward Elgar Publishing Limited.

Loke, Y.K., C.S. Kwok, and S. Singh (2011). Comparative cardiovascular effects of thiazolidinediones: a systematic review and meta-analysis of observational studies. *British Medical Journal* 342: d1309.

Longworth, L., J. Youn, L. Bojke, S. Palmer, S. Griffin, E. Spackman, and K. Claxton (2013). When does NICE recommend the use of health technologies within a programme of evidence development? A systematic review of NICE guidance. *Pharmacoeconomics* 31(2): 137-149.

Lubitz, J. D. and G. F. Riley (1993). Trends in Medicare payments in the last year of life. *New England Journal of Medicine* 328(15): 1092–1096.

Luce, B.R. and R.E. Brown (1995). The use of technology assessment by hospitals, health maintenance organizations, and third-party payers in the United States. *International Journal of Technology Assessment in Health Care* 11(1): 79-92.

Luce, B. R., and R. S. Cohen. (2009). Health technology assessment in the United States. *International Journal of Technology Assessment in Health Care* 25(suppl. 1): 33–41.

Luce, B. R., M. Drummond, B. Jonsson, P. J. Neumann, J. S. Schwartz, U. Siebert, and S.D. Sullivan. (2010). EBM, HTA, and CER: Clearing the confusion. *Milbank Quarterly* 88(2): 256–276.

Lundkvist, J. (2002). Pricing and reimbursement of drugs in Sweden. *European Journal of Health Economics* 3(1): 66-70.

Maciejewski, M.L., J.F. Farley, J. Parker, and D. Wansink (2010). Copayment reductions generate greater medication adherence in targeted patients. *Health Affairs* 29: 2002-2008.

Maio, V., and J. J. Gagne. (2010). Impact of ALLHAT publication on antihypertensive prescribing patterns in Regione Emilia-Romagna, Italy. *Journal of Clinical Pharmaceutical Therapy* 35: 55–61.

Majone, G. (1996). *Regulating Europe*. London/New York: Routledge.

Majone, G. (1997). The agency model: The growth of regulation and regulatory institutions in the European Union. *EIPASCOPE* 3: 1-6.

Majone, G. (2001). Non majoritarian institutions and the limits of democratic governance: A political transaction-cost approach. *Journal of Institutional and Theoretical Governance* 157: 57-78.

Malakoff, D. (2001). Memo to congress: Get better advice. *Science* 292(5525): 2229–2230.

Malenka, D.J., A.V. Kaplan, S.M. Sharp, and J.E. Wennberg (2005). Post-marketing surveillance of medical devices using Medicare claims. *Health Affairs* 24(4): 928-937.

Malin, J. L. (2010). Wrestling with the high price of cancer care: should we control costs by individuals' ability to pay or society's willing to pay? *Journal of Clinical Oncology* 28(20): 3212–3214.

Manchikanti, L., F.J. Falco, M.V. Boswell, and J.A. Hirsch (2010). Facts, fallacies, and politics of comparative effectiveness research: Part 1. Basic considerations. *Pain Physician* 13(1): E23-E54.

Mascini, P. and E. Van Wijk (2009). Responsive regulation at the Dutch food and consumer product safety authority: an empirical assessment of assumptions underlying the theory. *Regulation & Governance* 3: 27-47.

Mason, A. and M. Drummond (2009). Public funding of new cancer drugs: Is NICE getting nastier? *European Journal of Cancer* 45(7): 1188-1192.

Mason, A., M. Drummond, S. Ramsey, J. Campbell, and D. Raisch (2010). Comparison of anticancer drug coverage decisions in the United States and the United Kingdom: Does the evidence support the rhetoric? *Journal of Clinical Oncology* 28: 3234-3238.

Maynard, A. and D. McDaid (2003). Evaluating health interventions. Exploiting the potential. *Health Policy* 63: 215-226.

McCabe, C., K. Claxton, and A.J. Culyer (2008). The NICE cost-effectiveness threshold: what it is and what it means. *Pharmacoeconomics* 26(9): 733-744.

McGlynn, E. A., and R. H. Brook (2001). Keeping quality on the policy agenda. *Health Affairs* 20(3): 82-90.

McGregor, M. and J.M. Brophy (2005). End-user involvement in health technology assessment (HTA) development: A way to increase impact. *International Journal of Technology Assessment in Health Care* 21: 263-267.

McNeil, C. (2007). Sticker shock sharpens focus on biologics. *Journal of National Cancer Institute* 99(12): 910-914.

Medical Device Amendments Pub L No. 94-295, 90 Stat 539(1976).

Medpac (2010). *Report to Congress: Medicare Payment Policy*. Washington, DC: Medpac.

Medpac. (2008). *Report to Congress: Reforming the Delivery System*. Washington, DC: Medpac.

Meier, B. (2013). FDA to tighten regulation of all-metal hip implants. *New York Times*, January 16, 2013.

Melander, H., J. Ahlqvist-Rastad, G. Meijer, and B. Beerman (2003). Evidence b(i)ased medicine – selective reporting from studies sponsored by pharmaceutical industry: Review of studies in new drug applications. *British Medical Journal* 326: 1171.

Menon, D. and L.A. Topfer (2000). Health technology assessment in Canada. A decade in review. *International Journal of Technology Assessment in Health Care* 16(3): 896-902.

Meropol, N. J. and K. A. Schulman (2009). Cost of cancer care: issues and implications. *Journal of Clinical Oncology* 25(2): 180–186.

Meropol, N. J., D. Schrag, T. J. Smith, T. M. Mulvey, R.M. Langdon, D. Blum, P. A. Ubel and L. E. Schnipper (2009). American Society of Clinical Oncology guidance statement: the cost of cancer care. *Journal of Clinical Oncology* 27(23): 3868–3874.

Messner, D.A. and Tunis, S.R. (2012). Current and future state of FDA-CMS parallel reviews. *Clinical Pharmacology and Therapeutics* 91: 383-385.

Miles, M.B., A.M. Huberman, and J. Saldana (2014). *Qualitative Data Analysis: A Methods Sourcebook*. Third edition. London: Sage.

Milewa, T. and C. Barry (2005). Health policy and the politics of evidence. *Social Policy and Administration* 39(3): 498–412.

Milewa, T. (2006). Health technology adoption and the politics of governance in the UK. *Social Science and Medicine* 63: 3102–3112.

Miller, R. (2011). Debate on 510(k)s and recalls reaches Capitol Hill. *Medscape*. February 18. www.medscape.com/viewarticle/737721.

Miller, F.G. and S.D. Pearson (2008). Coverage with evidence development: Ethical issues and policy implications. *Medical Care* 46: 746-751.

Minghetti, P., P. Rocco, L. Del Vecchio, and F. Locatelli (2011). Biosimilars and regulatory authorities. *Nephron Clin Pract* 117(1): 1-7.

Moat, M., P. Ludman, M. Belder, et al. (2011). Long-term outcomes after transcatheter aortic valve implantation in high-risk patients with severe aortic stenosis. The U.K. TAVI (United Kingdom Transcatheter Aortic Valve Implantation) registry. *Journal of the American College of Cardiology* 58(20): 2130-38.

Mohr, P., S. Tunis, R. Sabharwal, et al. (2010). *The Comparative Effectiveness Research Landscape in the United States and its Relevance to the Medicare Program*. Baltimore, MD: Center for Medical Technology Policy.

Mohr, P. and S. Tunis (2010). Access with evidence development. *Pharmacoeconomics* 28: 153-162.

Moise, P. and E. Docteur (2007). *Pharmaceutical Pricing and Reimbursement Policies in Sweden. OECD Health Working Paper*. Paris: OECD.

Montoya, M.A. and F. Trillas (2009). The measurement of regulator independence in practice: Latin America and the Caribbean. *International Journal of Public Policy* 4(1-2): 113-134.

Morgan, S., M. McMahon, C. Mitton, E. Roughead, P. Kanavos, and D. Menon (2006). Centralized drug review processes in Australia, New Zealand, the United Kingdom, and Canada. *Health Affairs* 25(2): 337-347.

Mossialos, E., M. Mrazek, and T. Walley (2004). *Regulating Pharmaceuticals in Europe: Striving for Efficiency, Equity, and Quality*. Maidenhead: Open University Press.

Mossialos, E. and A. Oliver (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(4): 291-306.

Mossialos, E., D. Brogan, and T. Walley (2006). Pharmaceutical policy in Europe: Weighing up the options. *International Social Security Review* 59(3): 3-25.

Motola, D., F. De Ponti, E. Poluzzi, N. Martini, P. Rossi, M.C. Silvani, A. Vaccheri, and N. Montanaro (2006). An update on the first decade of the European centralized procedure: How many innovative drugs? *British Journal Clin Pharmacol* 62(5): 610-616.

Mukherjee, S. (2010). *The Emperor of All Maladies: A Biography of Cancer*. New York: Scribner.

Mulgan, R. (2000). Accountability: An ever-expanding concept? *Public Administration* 56: 25-36.

Mullins, C.D., R. Montgomery, and S. Tunis (2010). Uncertainty in assessing value of oncology treatments. *Oncologist* 15(Supplement 1): 58–64.

Munos, B. (2009). Lessons for 60 years of pharmaceutical innovation. *Nature Reviews* 8: 959-968.

Naci, H., J.D. Cylus, A. Sato, S. Vondoros, and K. Perampaladas (2012). Raising the bar for market authorization. *BMJ* 344: e4261.

Nadler E., B. Eckert, and P.J. Neumann (2006). Do oncologists believe new cancer drugs offer good value? *Oncologist* 11(2): 90–95.

Nasser, M. and P. Sawicki (2009). *Institute for Quality and Efficiency in Health Care: Germany*. New York: The Commonwealth Fund.

National Cancer Director (2006). *Usage of Cancer Drugs Approved by NICE*. London: Department of Health.

National Health Authority (2007). *Service to the Community (SERC): How to Take into Account other Dimensions Other than Medical in Practice?* Paris: National Health Authority.

National Institute for Health and Clinical Excellence (NICE). (2002). *Guidance on the Use of Metal-On-Metal Hip Resurfacing Arthroplasty (TA44)*. London: NICE.

National Institute for Health and Clinical Excellence (NICE) (2006a). *NICE Implementation Uptake Report: Drugs for Obesity*. London: NICE.
<http://www.nice.org.uk/page.aspx?o5390757>

National Institute for Health and Clinical Excellence (NICE) (2006b). *NICE Implementation Uptake Report: Cox II Selective Inhibitors*. London: NICE.
<http://www.nice.org.uk/page.aspx?o5390760>

National Institute for Health and Clinical Excellence (NICE). (2006c). *Laparoscopic Surgery for the Treatment of Colorectal Cancer (TA105)*. London: NICE.

National Institute for Health and Clinical Excellence (NICE) (2008a). *Guide to the Methods of Technology Appraisal*. London: NICE.

National Institute for Health and Clinical Excellence (NICE) (2008b). *Social Value Judgments: Principles for the Development of NICE Guidance, 2nd edition*. London: NICE.

National Institute for Health and Clinical Excellence (NICE) (2009). *Net Cost Estimate, Implementation Programme*. London: NICE.

National Institute for Health and Clinical Excellence (NICE). (2011). *MIST Therapy System for the Promotion of Wound Healing in Chronic and Acute Wounds (MTG5)*. London: NICE.

Nicod, E. and P. Kanavos (2012). Commonalities and differences in HTA outcomes: A comparative analysis of five countries and implications for coverage decisions. *Health Policy* 108(2-3): 167-177.

Lee, J.L., M.A. Fisher, W.H. Shrank, J.M. Polinski, and N.K. Choudhry (2012). A systematic review of reference pricing: Implications for US prescription drug spending. *American Journal of Managed Care* 18(11): e429-e437.

Lehoux, P. and B. Williams-Jones (2007). Mapping the integration of social and ethical issues in health technology assessment. *International Journal of Technology Assessment in Health Care* 23(1): 9-16.

Neumann, P.J., E.A. Sandberg, C.M. Bell, P.W. Stone and R.H. Chapman (2000). Are pharmaceuticals cost-effective? A review of the evidence. *Health Affairs* 19(2): 92–109.

Neumann, P.J. (2004). *Using Cost-Effectiveness Analysis to Improve Health Care: Opportunities and Barriers*. New York: Oxford University Press.

Neumann, P.J., A B. Rosen, and M.C. Weinstein (2005). Medicare and cost-effectiveness analysis. *New England Journal of Medicine* 353(14): 1516–1522.

Neumann, P.J., N. Nivi, M.T. Beinfeld, B. Levine, P.S. Keenan, E.K. Halpern, and G.S. Gazelle. (2005). Medicare's national coverage decisions, 1999–2003: Quality of evidence and review times. *Health Affairs* 24(1): 243–254.

Neumann, P.J. (2006). Emerging lessons from the drug effectiveness review project. *Health Affairs* 24: 80–92.

Neumann, P.J., M.S. Kamae, and J.A. Palmer (2008). Medicare's national coverage decisions for technologies, 1999-2007. *Health Affairs* 27(6): 620-631.

Neumann, P.J. (2010). Legislating against use of cost-effectiveness information. *New England Journal of Medicine* 363(16): 1495-1497.

Neumann, P.J., J.A. Palmer, E. Nadler, C. Fang, and P. Ubel (2010). Cancer therapy costs influence treatment: a national survey of oncologists. *Health Affairs* 29(1): 196-202.

Neyt, M., H. Van Brabandt, S. Van De Sande, and S. Devriese (2011). *Health Technology Assessment. Transcatheter Aortic Valve Implantation (TAVI): A Health Technology Assessment Update*. Belgian Health Care Knowledge (KCE) Reports 163C. D/2011/10.273/48. Brussels: KCE.

Nielsen, V.L. and C. Parker (2009). Testing responsive regulation in regulatory enforcement. *Regulation & Governance* 3: 376-99.

Nissan, S. (2010). *Rosiglitazone: A Critical Review*.
www.fda.gov/downloads/AdvisoryCommittees/CommitteeMeetingMaterials/Drugs/EndocrinologicalandMetabolicDrugsAdvisoryCommittee/UCM218483.pdf.

Littlejohns, P., A. Weale, K. Chalkidou, R. Faden, and Y. Teerawattananon (2012). Social values and health policy: A new international research programme. *Journal of Health Organization and Management* 26(3): 285-292.

Noorani, H.Z., D.R. Husereau, R. Boudreau, and B. Skidmore (2007). Priority setting for health technology assessments: A systematic review of current practical approaches. *International Journal of Technology Assessment in Health Care* 23(3): 310-315.

Obama-Biden Presidential Campaign (2008). Barack Obama and Joe Biden's plan to lower Health care costs and ensure affordable, accessible health care coverage for all. Washington, DC: Obama-Biden Presidential Campaign, Oct. 3.

Oberlander, J., T. Marmor, and L. Jacobs (2001). Rationing medical care: Rhetoric and reality in the Oregon health plan. *Canadian Medical Association Journal* 164(1): 1583–1587.

Oberlander, J. (2007). Health reform interrupted: The unraveling of the Oregon health plan. *Health Affairs* 26(1): w96-w105.

O’Conner, A.B. (2010). Building comparative efficacy and tolerability into the FDA approval process. *JAMA* 303(10): 979-980.

Oortwijn, W.J., S.R. Hanney, A. Ligtoet, et al. (2008). Assessing the impact of health technology assessment in the Netherlands. *International Journal of Technology Assessment in Health Care* 24(3): 259-269.

Oortwijn, W.J., J. Mathijssen, and D. Banta (2010). The role of health technology assessment on pharmaceutical reimbursement in selected middle-income countries. *Health Policy* 95(2-3): 174-184.

Organization for Economic Cooperation and Development (OECD) (1997). *The OECD Report on Regulatory Reform: Synthesis Report*. Paris: OECD.

Organization for Economic Cooperation and Development (OECD) (1999). *Results Based Management in the Development of Co-Operation Agencies: A Review of Experience*. Paris: OECD.

Organization for Economic Cooperation and Development (OECD) (2001). *Citizens as Partners: Information, Consultation and Public Participation in Policy-Making*. Paris: OECD.

Organization for Economic Cooperation and Development (OECD) (2003). *Special Group on Regulatory Policy: Independent Regulators, Political Challenges and Institutional Design – Discussion note*. SG/SGR(2003)3; 24 February. Paris: OECD.

Organization for Economic Cooperation and Development (OECD) (2004). *Regulatory Performance: Ex Post Evaluation of Regulatory Policies*. Paris: OECD.

Organization for Economic Cooperation and Development (OECD) (2011). *OECD Health Data 2011*. Paris: OECD.

Orvain, J., B. Xerri, and Y. Matillon (2004). Overview of health technology assessment in France. *International Journal of Technology Assessment in Health Care* 20(1): 25-34.

Office of Technology Assessment (OTA) (1996). *The OTA legacy*, Washington, DC: Government Printing Office. <http://www.princeton.edu/ota/>

Ojo, M. (2009). *Response Regulation: Achieving the Right Balance between Persuasion and Penalisation*. MPRA Paper 15543, University Library of Munich, Germany.

Oliver, A., E. Mossialos, and R. Robinson (2004). Health technology assessment and its influence on health-care priority setting. *International Journal of Technology Assessment in Health Care* 20(1): 1–10.

Oliver, A. (2007). The Veterans Health Administration: An American success story? *Milbank Quarterly* 85(1): 5–35.

Oliver, A. and C. Sorenson (2009). The limitations and challenges to the economic evaluation of health technologies. *The Economics of New Health Technologies: Incentives, Organization, and Financing*. J. Costa-Font, C. Courbage and A. McGuire, eds. Oxford: Oxford University Press.

Olson, M. (1965). *The Logic of Collective Action*. Cambridge: Harvard University Press.

- Orzack, L. (1996). Professionals, consumers and the European Medicines Agency: Policy-making in the European Union. *Current Research on Occupations and Professionals* 9: 9-29.
- OTA (Office of Technology Assessment). 1994. *Identifying Health Technologies That Work: Searching for Evidence*. Washington, DC: OTA.
- OTA (Office of Technology Assessment). 1995. *Health Care Technology and Its Assessment in Eight Countries*. OTA-BP-H-140. Washington, DC: OTA.
- Owen-Smith, A., J. Coast and J. Donovan (2010). The usefulness of NICE guidance in practice: different perspectives of managers, clinicians, and patients. *International Journal of Technology Assessment in Health Care* 26(3): 317–322.
- Palmer, S. and P. Smith (2000). Incorporating option values into the economic evaluation of health care technologies. *Journal of Health Economics* 19: 755-766.
- Paris, V. and E. Docteur (2006). *Pharmaceutical Pricing and Reimbursement Policies in Canada*. Paris: Organisation for Economic Co-operation and Development.
- Paris, V. and E. Docteur (2007). *Pharmaceutical Pricing and Reimbursement Policies in Sweden*. Paris: Organisation for Economic Co-operation and Development.
- Patel, K. (2010). Health reform's tortuous route to the Patient-Centered Outcomes Research Institute. *Health Affairs* 29(10): 1777–1782.
- Patel, P., S. Roberts, K. Sparrowhawk and J. Ng Haing (2010). Evaluating comparative effectiveness research for oncology in the US: An analysis based on lessons from health technology assessment in Europe. ISPOR 15th Annual International Meeting, May 15–19th, Atlanta, GA.
- Payne, G. and J. Payne (2004). *Key Concepts in Social Research*. London: Sage.

Pear, R. (1991). Medicare to weigh cost as a factor for reimbursement. *New York Times*, April 21.

Pearson, S.D. and P.B. Bach (2010). How Medicare could use comparative effectiveness research in deciding on new coverage and reimbursement. *Health Affairs* 29(10): 1796–1804.

Pedersen, K. (2003). Pricing and reimbursement of drugs in Denmark. *European Journal of Health Economics* 4(1): 60-65.

Perry, S., E. Gardner, and M. Thamer (1997). The status of health technology assessment worldwide: Results of an international survey. *International Journal of Technology Assessment in Health Care* 13(1): 81-98.

Perry, S. and M. Tharner (1999). Medical innovation and the critical role of health technology assessment. *Journal of the American Medical Association* 282: 1869-1872.

Peltzman, S. (1998). *Political Participation and Government Regulation*. Chicago: University of Chicago Press.

Perehudoff, S.K. and T.L. Alves (2011). *The Patient and Consumer Voice and Pharmaceutical Industry Sponsorship*. Amsterdam: Health Action International Europe.

<http://www.haiweb.org/31012011/31%20Jan%202011%20HAI%20EUROPE%20Research%20Article%20Patient%20&%20consumer%20voice%20and%20pharmaceutical%20industry%20sponsorship.pdf>

Perleth, M., B. Gibis and B. Gohlen (2009). A short history of health technology assessment in Germany. *International Journal of Technology Assessment in Health Care* 25(S1): 112–119.

- Permanand, G. and E. Mossialos (2005). Constitutional asymmetry and pharmaceutical policy-making in the European Union. *Journal of European Public Policy* 12(4): 687-709.
- Permanand, G. (2006). *EU Pharmaceutical Regulation: the Politics of Policy-Making*. European Policy Research Unit Series. Manchester, UK: Manchester University Press.
- Permanand, G., E. Mossialos, and M. McKee (2006). Regulating medicines in Europe: The European Medicines Agency, marketing authorisation, transparency and pharmacovigilance. *Clinical Medicine* 6(1): 87-90.
- Permanand, G., E. Vos (2010). EU regulatory agencies and health protection. *Health Systems Governance in Europe*. E. Mossialos, G. Permanand, R. Baeten, and T.K. Hervey, eds. Cambridge: Cambridge University Press.
- Persson, U. and J. Hjelmgren (2003). Health services need knowledge of how the public values health. *Lakartidningen* 100(43): 3436–3437.
- Persson, U. (2012) *Value-Based Pricing in Sweden: Lessons For Design?* OHE Seminar Briefing, November. London: Office of Health Economics.
- Phillips, L. (2011). EU drugs regulator accused of being too cozy with Big Pharma. *Euobserver*, November 5.
- Phillips, Z., L. Bojke, M. Sculpher, K. Claxton, and S. Golder (2006). Good practice guidelines for decision-analytic modelling in health technology assessment: A review and consolidation of quality assessment. *Pharmacoeconomics* 24(4): 355-371.
- Powell, W.W. and P.J. DiMaggio (1991). *The New Institutionalism in Organizational Analysis*. Chicago: Chicago University Press.
- PPRC (Physician Payment Review Commission) (1995). *Annual Report to Congress*.

Washington, DC: PPRC.

PricewaterhouseCoopers (PwC) (2011). *Medical Technology Innovation Scorecard*. New York: PwC.

Punch, K.F. (2005). *Introduction to Social Research: Quantitative and Qualitative Research*. Second Edition. London: Sage.

Radaelli, C.M. (2000). *Whither Europeanization? Concept Stretching and Substantive Change*. European Integration Online Papers (EIoP) Vol. 4.

Radaelli, C.M. and V.A. Schmidt (2004). Discourse and policy change in Europe: Conceptual and methodological issues. *West European Politics* 27(2): 183-210.

Radaelli, C.M. and F. DeFrancesco (2007). Indicators of regulatory quality. *Regulatory Impact Assessment: Towards Better Regulation?* C. Kirkpatrick and D. Parker, eds. Cheltenham: Edward Elgar.

Raftery, J. (2010a). Should NICE's threshold range for cost per QALY be raised? No. *British Medical Journal* 338: b185.

Raftery, J. (2010b). Multiple sclerosis risk sharing scheme: a costly failure. *British Medical Journal* 340: c1672.

Regnstrom, J., F. Koenig, B. Aronsson, T. Reimer, K. Svendsen, S. Tsigkos, B. Flamion, H.G. Eichler, and S. Vamvakas (2010). Factors associated with success of market authorisation applications for pharmaceutical drugs submitted to the European Medicines Agency. *Eur J Clin Pharmacol* 66(1): 39-48.

Resnic, F.S. and S.L. Normand (2012). Postmarketing surveillance of medical devices – filling in the gaps. *New England Journal of Medicine* 366(10): 875-877.

Ramsey, S.D. and S.D. Sullivan SD (2005). Evidence, economics, and emphysema: Medicare's long journey with lung volume reduction surgery. *Health Affairs* 24(1): 55-66.

Ramsey, S.D. and A. Schickedanz (2010). How should we define value in cancer care? *The Oncologist* 15(Supplement 1): 1-4.

Rawlins, M. D. and K. Chalkidou (2011). The opportunity cost of cancer care: a statement from NICE. *Lancet Oncology* 12(10): 931-932.

Rawlins, M. and A. Culyer (2004). National Institute for Clinical Excellence and its value judgements. *British Medical Journal* 329: 224.

Redberg, R. F. and J. Walsh (2008). Pay now, benefits may follow: The case of cardiac computed tomographic angiography. *New England Journal of Medicine* 359(22): 2309-2311.

Regnstrom, J., F. Koenig, B. Aronsson, et al. (2010). Factors associated with success of market authorisation applications for pharmaceutical drugs submitted to the European Medicines Agency. *Eur J Clin Pharmacol* 66(1): 39-48.

Remuzat, C., M. Toumi, and B. Falissard (2013). New drug regulations in France: What are the impacts on market access? Part 1 – Overview of new drug regulations in France. *Journal of Market Access and Health Policy* 1: 20891.

Rinta, S. (2001). Pharmaceutical pricing and reimbursement in Finland. *European Journal of Health Economics* 2(3): 128-135.

Rising, K., P. Bacchetti, and L. Bero (2008). Reporting bias in drug trials submitted to the Food and Drug Administration: Review of publication and presentation. *PLoS Medicine* 5 (11): e217.

Rittberger B, Wonka A. (2011). Introduction: Agency governance in the European Union. *Journal of European Public Policy*, 18(6): 780-89.

- Robinson, J.C. (2010). Comparative effectiveness research: From clinical information to economic incentives. *Health Affairs* 29(10): 1788–1795.
- Rochaix, R. and B. Xerri (2009). *National Authority for Health: France*. New York: The Commonwealth Fund.
- Roughead, E.E., R. Lopert, and L.N. Sansom (2007). Prices for innovative pharmaceutical products that provide health gain: A comparison between Australia and the United States. *Value in Health* 10(6): 514-520.
- Rovira, J. and J. Darba (2001). Pharmaceutical pricing and reimbursement in Spain. *European Journal of Health Economics* 2(1): 39-43.
- Scharpf, F. (1999). *Governing in Europe*. Oxford: Oxford University Press.
- Scheller-Kreinsen, D., W. Quentin, and R. Busse (2012). DRG-based hospital payment systems and technological innovation in 12 European countries. *Value in Health* 14(8): 1166-1172.
- Schlozman, K.L. and J.T. Tierney (1986). *Organized Interests and American Democracy*. New York: Harper and Row.
- Schluessmann, E., P. Diel, E. Aghayev, T. Zweig, P. Moulin, C. Roder, and the SWISSspine Registry Group (2009). SWISSspine: A nationwide registry for health technology assessment of lumbar disc prostheses. *Eur Spine Journal* 18(6): 851-861.
- Schluessmann, E., E. Aghayev, L. Staub, P. Moulin, T. Zweig, C. Roder, and the SWISSspine Registry Group (2010). SWISSspine: The case of a governmentally required HTA-registry for total disc arthroplasty. Results of cervical disc prostheses. *Spine* 35(24): E1397-E1405.
- Schnipper, L.E., N.J. Meropol, and D.W. Brock (2010). Value and cancer care: toward an equitable future. *Clinical Cancer Research* 16(24): 6004–6008.

Schmidt, H. and J. Kreis (2009). Lessons from abroad. *Hastings Center Report* 39(6): 20–22.

Schoen, C., S. Guterman, A. Shih, J. Lau, S. Kasimow, A. Gauthier, and K. Davis. (2007). *Bending the Curve: Options for Achieving Savings and Improving Value in U.S. Health Spending*. New York: Commonwealth Fund.

Schrag, S. (2004). The price tag on progress – chemotherapy for colorectal cancer. *New England Journal of Medicine* 351(4): 317–319.

Schrag, S. and M. Hanger (2007). Medical oncologists views on communicating with patients about chemotherapy costs: A pilot survey. *Journal of Clinical Oncology* 25: 233–237.

Schickedanz, A. (2010). Of value: A discussion of cost, communication, and evidence to improve cancer care. *The Oncologist* 15(Supplement 1): 73–79.

Schreyogg, J., M. Baulmer, and R. Busse (2009). Balancing adoption and affordability of medical devices in Europe. *Health Policy* 92: 218-224.

Scott, C. (2000). Accountability in the regulatory state. *Journal of Law and Society* 27: 38-60.

Scott, J. (1990). *A Matter of Record: Documentary Sources in Social Research*. Cambridge: Polity Press.

Sedrakyan, A., S.T. Normand, S. Dabic, S., et al. (2011). Comparative assessment of implantable hip devices with different bearing surfaces: Systematic appraisal of evidence. *BMJ* 343: d7434.

Sheldon, T.A., N. Cullum, P. Dawson, A. Lankshear, K. Lawson, I. Watt, P. West, D. Wright and J. Wright (2004). What’s the evidence that NICE guidance has been

implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews. *British Medical Journal* 329: 999.

Shih, Y.C. and M.T. Halpern (2008). Economic evaluations of medical care interventions for cancer patients: how, why, and what does it mean? *CA: A Cancer Journal for Clinicians* 58(4): 231–244.

Shuren, J. (2011). Debate on 510(k)s and recalls reaches Capitol Hill. *Theheart.org*, February 18.

Siebert, M., L.C. Clauss, M. Carlisle, et al. (2002). Health technology assessment for medical devices in Europe: What must be considered. *International Journal of Technology Assessment in Health Care* 18(3): 733-740.

Sigmund, H. and F.B. Kristensen. (2002). Does health technology assessment benefit health services and politics? The experiences of an established HTA institution: the Danish Centre for Evaluation and HTA. *European Journal of Health Economics* 3(1): 54-58.

Smith, T. J. and B. E. Hillner (2010). Concrete options and ideas for increasing value in oncology care: the view from one trench. *Oncologist* 15(Supplement 1): 65–72.

Smith, T.J. and B.E. Hillner (2011). Bending the cost curve in cancer care. *New England Journal of Medicine* 364(21): 2060–2065.

Slevin, M.L., L. Stubbs, J.J. Plant, P. Wilson, W.M. Gregory, P.J. Armes and S.M. Downer (1990). Attitudes to chemotherapy: comparing views of patients with cancer with those of doctors, nurses, and general public. *British Medical Journal* 300: 1458–1460.

Soares, M. (2005). Off-label indications for oncology drug use and drug compendia: history and current status. *Journal of Oncology Practice* 1(3): 102–105.

Sorenson, C., M. Drummond, and P. Kanavos (2008a). *Ensuring Value for Money in Health Care: The Role of HTA in the European Union*. Copenhagen: European Observatory on Health Systems and Policies.

Sorenson, C., M. Drummond, P. Kanavos, and A. McGuire (2008b). *National Institute for Health and Clinical Excellence (NICE): How Does It Work and What Are the Implications for the US?* Reston, VA: National Pharmaceutical Council.

Sorenson, C., P. Kanavos, and M. Karamalis (2009). HTA in central and eastern Europe: Current status, challenges and opportunities. *Journal of Medical Device Regulation* 6(1): 34-45.

Sorenson, C. (2010). *Use of Comparative Effectiveness Research in Drug Coverage and Pricing Decisions: A Six-Country Comparison*. New York: The Commonwealth Fund.

Sorenson, C. and P. Kanavos (2011). Medical technology procurement in Europe: A cross-country comparison of current practice and policy. *Health Policy* 100(1): 43-50.

Sorenson, C., R. Tarricone, M. Siebert, and M. Drummond (2011a). Applying health economics for policy decision making: Do devices differ from drugs? *Europace* 13(suppl 2): ii54-ii58.

Sorenson, C., H. Naci, J. Cylus, and E. Mossialos (2011b). Evidence of comparative efficacy should have a formal role in European drug approvals. *BMJ* 343: d4849.

Sorenson, C., M. Drummond, and R.L. Burns (2013). Evolving reimbursement and pricing policies for devices in Europe and the United States should encourage greater value. *Health Affairs* 32(4): 788-796.

Sox, H. (2009). Evaluating off-label uses of anticancer drugs: time for a change. *Annals of Internal Medicine* 150(5): 353–354.

- Spaulding, A., A.M. Fendrick, W.H. Herman, et al. (2009). A controlled trial of value-based insurance design – The MHealthy: Focus on diabetes (FOD) trial. *Implementation Science* 4:19. doi:10.1186/1748-5908-4-19.
- Squires, D. (2012). *Explaining High Health Care Spending in the United States: An International Comparison of Supply, Utilization, Prices, and Quality*. New York (NY): Commonwealth Fund.
- Stafinski, T., D. Menon, C. Davis, and C. McCabe (2011a). Role of centralized review processes for making reimbursement decisions on new health technologies in Europe. *Clinicoecon Outcomes Research* 3: 117-186.
- Stafinski, T., D. Menon, D.J. Philippon, et al. (2011b). Health technology funding decision-making processes around the world: the same, yet different. *Pharmacoeconomics* 29: 475-495.
- Steinberg, E.P. and B. Luce (2005). Evidence Based? Caveat Emptor! *Health Affairs* 24: 80–92.
- Stern, J.M. and R.J. Simes (1997). Publication bias: evidence of delayed publication in a cohort study of clinical research protocols. *BMJ* 315(7109): 640-645.
- Stern, J.M. (1997). What makes an independent regulator independent? *Business Strategy Review* 8(2): 67-74.
- Stevens, A., R. Milne, and A. Burls (2003). Health technology assessment: History and demand. *Journal of Public Health Medicine* 25(2): 98-101.
- Stevens, A. and R. Milner (2004). Health technology assessment in England and Wales. *International Journal of Technology Assessment in Health Care* 20(1): 11-24.
- Stewart, D.W. and M.A. Kamins (1999). *Secondary Research: Information Sources and Methods*. Second Edition. London: Sage.

Straus, S.E. (2004). What has evidence based medicine does for us?' *British Medical Journal* 329: 987–988.

Sullivan, R., J. Peppercorn, K. Sikora, et al. (2011). Delivering affordable cancer care in high-income countries. *Lancet Oncology* 12(10): 933–974.

Sullivan, S.D., J. Watkins, B. Sweet, and S.D. Ramsey (2009). Health technology assessment in health-care decisions in the United States.” *Value in Health* 12(suppl. 2): S39–S44.

Sulmasy, D.P. (2007). Cancer care, money, and the value of life: whose justice? Which rationality? *Journal of Clinical Oncology* 25(2): 217–222.

Swedish Dental and Pharmaceutical Benefits Agency (2010). *Developing the Pricing System for Pharmaceuticals Could Save Billions*. <http://www.tlv.se/in-english/news/developingthe-pricing-system-for-pharmaceuticals-could-save-billions/>.

Sweet, B.V., A.K. Schwemm, and D.M. Parsons (2011). Review of the processes for FDA oversight of drugs, medical devices, and combination products. *J Manage Care Pharm* 17: 40-50.

Szabo, L. (2008). Study: Many cancer patients forgoing care because of cost. *USA Today*, October 13.

Syrett, K. (2003). A technocratic fix to the “legitimacy problem”? The Blair government and health care rationing in the United Kingdom. *Journal of Health Politics, Policy and Law* 28(4): 715–746.

Tangka, F.K., J.G. Trogon, L.C. Richardson, D. Howard, S.A. Sabatino and E.A. Finkelstein (2010). Cancer treatment cost in the United States. Has the burden shifted over time? *Cancer* 116(14): 3477–3484.

Taylor, R.S. and C.P. Iglesias (2009). Assessing the clinical and cost-effectiveness of medical devices and drugs: Are they that different? *Value in Health* 12(4): 404-406.

Temel, J.S., V.S. Jackson and J.A. Billings (2010). Early palliative care for patients with metastatic non-small-cell lung cancer. *New England Journal of Medicine* 36(8): 733–742.

Thatcher, M. (2002). Delegation to independent regulatory agencies: pressures, functions and contextual mediation. *West European Politics* 25(1): 125-147.

Thatcher, M. (2011). The creation of European regulatory agencies and its limits: A comparative analysis of European delegation. *Journal of European Public Policy* 18(6): 780-809.

Thomson, S. and E. Mossialos (2010). *Primary Care and Prescription Drugs: Coverage, Cost-Sharing, and Financial Protection in Six European Countries*. New York: The Commonwealth Fund.

Tol, J., M. Koopman, A. Cats, et al. (2009). Chemotherapy, bevacizumab and cetuximab in metastatic colorectal cancer. *New England Journal of Medicine* 360: 563–572.

Torbica, A. and G. Cappellaro (2010). Uptake and diffusion of medical technology innovation in Europe: What role for funding and procurement policies? *Journal of Medical Marketing* 10(1): 61-69.

Towse, A. (2009). Should NICE's threshold range for cost per QALY be raised? Yes. *British Medical Journal* 338: b181.

Towse, A. and L.P. Garrison (2010). Can't get no satisfaction? Will pay for performance help? Toward an economic framework for understanding performance-based risk-sharing agreements for innovative medical products. *Pharmacoeconomics* 28(2): 93-102.

Trondal, J. and L. Jeppesen (2008). Images of agency governance in the European Union. *West European Politics* 31(3): 417-41.

Trubek, D.M. and L.G. Trubeck (2005). Hard and soft law in the construction of social Europe: the role of the open method of co-ordination. *European Law Journal* 11(3): 343-364.

Trueman, P., D.L. Grainger, and K.E. Downs (2010). Coverage with evidence development: Applications and issues. *International Journal of Technology Assessment in Health Care* 26: 79-85.

Tu, J.V., J. Bowen, M. Chiu, et al. (2007). Effectiveness and safety of drug-eluting stents in Ontario. *New England Journal of Medicine* 357(14): 1393–1402.

Turner, E.H., A.M. Matthews, E. Lindardatos, R.A. Tell, and R. Rosenthal (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine* 358: 252-260.

Tunis, S. R., and H. Gelbrand. (1994). “Health Care Technology in the United States.” *Health Policy* 30: 335–96.

Tunis, S.R. and S.D. Pearson (2006). Coverage options for promising technologies: Medicare’s coverage with evidence development. *Health Affairs* 25(5): 1218–1230.

Tunis, S., R.A. Berenson, S.E. Phurrough, and P.E. Mohr (2011). *Improving the Quality and Efficiency of the Medicare Program through Coverage Policy*. Washington, DC: Urban Institute.

Ubel, P., C. Jepson, J. Baron, J. C. Hershey and D.A. Asch (2003). The influence of cost-effectiveness information on physicians’ cancer screening recommendations. *Social Science and Medicine* 58(8): 1727–1736.

UK Better Regulation Task Force (2000). *Principles of Good Regulation*. London: Cabinet Office.

US Food and Drug Administration (FDA) (2005). *US FDA Center for Devices and Radiological Health Medical Devices Advisory Committee Circulatory Systems Devices Panel*. Silver Spring, MD: FDA.

US Food and Drug Administration (FDA) (2008). *Modifications to Devices Subject to Premarket Approval (PMA) - the PMA Supplement Decision*. Silver Spring, MD: FDA.

US Food and Drug Administration (FDA) (2011a). *A Delicate Balance: FDA and the Reform of the Medical Device Approval Process*. Testimony before Senate Special Committee on Agency, United States Senate, 112th Congress, First Session, April 13. www.fda.gov/NewsEvents/Testimony/ucm250709.htm.

US Food and Drug Administration (FDA) (2011b). *Understanding Barriers to Medical Device Quality*. Silver Spring, MD: FDA.

US Senate Committee on HELP (Health, Education, Labor, and Pensions) (2009). Enzi: No Washington Bureaucrat Should Deny or Delay Care for You and Your Family, June 22. http://www.help.senate.gov/old_site/Min_press/2009_06_22_a.pdf

Vedula, S.S., L. Bero, R.W. Scherer, and K. Dickersin (2009). Outcome reporting in industry-sponsored trials of Gabapentin for off-label use. *New England Journal of Medicine* 361: 1963-1971.

Velasco-Garrido, M. and R. Busse (2005). *Health Technology Assessment: An Introduction on Objectives, Role of Evidence, and Structure in Europe*. Brussels: European Observatory on Health Systems and Policies.

Velasco-Garrido, M., F.B. Kristensen, C.P. Nielsen and R. Busse (2008). *Health Technology Assessment and Health Policy-making in Europe: Current Status, Challenges and Potential*. Copenhagen: European Observatory on Health Systems and Policies.

- Van Brabandt, H., M. Neyt, and F. Hulstaert (2012). Transcatheter aortic valve implantation (TAVI): risky and costly. *British Medicines Journal* 345:e4710.
- Virgo, K.S., A.M. Vernava, W.E. Longo, M.E. Kissling, J.C. Lin, and F.E. Johnson (1995). Cost of patient follow-up after potentially curative colorectal cancer treatment. *Journal of the American Medical Association* 273: 1837–1841.
- Vogler, S., C. Habl, C. Leopold, I. Rosian-Schikuta, K. De Joncheere, and T.L. Thomsen (2009). *Pharmaceutical Pricing and Reimbursement Information (PPRI) Report*. Vienna, Austria: Gesundheit Osterreich GmbH.
- Vogler, S., N. Zimmerman, C. Leopold, and K. De Joncheere (2011). Pharmaceutical Policies in European Countries in Response to the Global Financial Crisis. *Southern Med Review* 4(2): 22-32.
- Vogler, S. (2012). The impact of pharmaceutical pricing and reimbursement policies on generic update: implementation of policy options on generics in 29 European countries – an overview. *Generics and Biosimilars Initiative Journal* 1(2): 93-100.
- Vos, E. (2002). European governance: Revitalizing the community method. *Towards a European Ius Commune in Legal Education and Research*. M. Faure, H. Schneider, and J. Smits, eds. Cambridge: Intersentia.
- Walker, S., M. Sculpher, K. Claxton, and S. Palmer (2012). *Coverage with Evidence Development, Only in Research, Risk Sharing or Patient Access Scheme? A Framework for Coverage Decisions*. CHE Research Paper 77. York: University of York, Centre for Health Economics.
- Weatherill, S.R. (2007). *Better Regulation*. Oxford: Hart Publishing.
- Wennberg, J.E. (1984). Dealing with medical practice variations: A proposal for action. *Health Affairs* 3(2): 6–32.

- Whicher, D.M., K. Chalkidou, I.A. Dhalla, L. Levin, and S. Tunis (2009). Comparative effectiveness research in Ontario, Canada: Producing relevant and timely information for health care decision makers. *Milbank Quarterly* 87(3): 585-606.
- Wiktorowicz, M.E. (2003). Emergent patterns in the regulation of pharmaceuticals: Institutions and interests in the United States, Canada, Britain and France. *Journal of Health Politics, Policy and Law* 28(4): 615-658.
- Wilensky, G.R. (2006). Developing a center for comparative effectiveness information. *Health Affairs* 25(6): 572–585.
- Wilensky, G.R. (2009). The policies and politics of creating a comparative clinical effectiveness research center. *Health Affairs* 28(4): 719-729.
- Will, G.F. (2009). How the GOP should measure the stimulus. *Washington Post*, January 29.
- Williams, I. and S. Bryan (2007). Understanding the limited impact of economic evaluation in health care resource allocation: a conceptual framework. *Health Policy* 80: 135–143.
- Williamson, R. (2010). Patient access schemes for high-cost cancer medicines. *Lancet* 11: 111-112.
- Wilmshurst P. The regulation of medical devices. *British Medicines Journal* 2011; 342: d2822.
- Wong, U., O. Hamilton, B. Egleston, K. Salador, C. Murphy, and N.J. Meropol (2010). Understanding how out-of-pocket expenses, treatment value, and patient characteristics influence treatment choices. *The Oncologist* 15: 566–576.

Woolf, S. H. and C. Henshall (2000). Health technology assessment in the United Kingdom. *International Journal of Technology Assessment in Health Care* 16(2): 591-625.

Wright, A.A., B. Zhang, A. Ray, J.W. Mack, E. Trice, T. Balboni, S.L. Mitchell, V.A. Jackson, S.D. Block, P.K. Maciejewski and H.G. Prigerson (2008). Associations between end-of-life discussions, patient mental health, medical care near death, and caregiver bereavement adjustment. *Journal of the American Medical Association* 300: 1665–1673.

Yfantopoulos, J. (2008). Pharmaceutical pricing and reimbursement reforms in Greece. *European Journal of Health Economics* 9(1): 87-97.

Zentner, A., M. Valasco-Garrido, and R. Busse (2005). Methods for the comparative evaluation of pharmaceuticals. *GMS Health Technology Assessment* 1: Doc09.

Zhang, B., A.A. Wright, H.A. Huskamp, M.E. Nilsson, M.L. Maciejewski, C.C. Earle, S. D. Block, P.K. Maciejewski and H.G. Prigerson (2009). Health care costs in the last week of life: associations with end-of-life conversations. *Archives of Internal Medicine* 169: 480–488.

Zhang, J. (2009). Push to compare treatments worries drug, device makers. *Wall Street Journal*, April 15.

Zuckerman, D.M., P. Brown, and S.E. Nissen (2011). Medical device recalls and the FDA approval process. *Arch Intern Med* 171(11): 1006-1011.

Appendices

Appendix A: List of publications reviewed

Adamski, J., B. Godman, and G. Ofierska-Sujkowska (2007). Risk sharing arrangements for pharmaceuticals: potential considerations and recommendations for European payers. *BMC Health Services Research* 10: 153.

American Society of Clinical Oncology (2007). Medicare's coverage with evidence development: A policy-making tool in evolution. *JOP* 3(6): 296-301.

Atkins, D., J. Siegel, and J. Slutsky (2005). Making policy when evidence is in dispute. *Health Affairs* 24(1): 102-113.

Briggs, A., K. Ritchie, E. Fenwick, K. Chalkidou, P. Littlejohns. (2010). Access with evidence development in the UK. Past experience, current initiatives and future potential. *Pharmacoeconomics* 28: 163-170.

Carlson, J.J., S.D. Sullivan, L.P. Garrison, P.J. Neumann, and D.L. Veenstr (2010). Linking payment and health outcomes: A systematic review and taxonomy of performance-based health outcomes agreements between healthcare payers and manufacturers. *Health Policy* 96(3): 179-190.

Carlson, J.J., L.P. Garrison, and S.D. Sullivan (2009). Paying for outcomes: innovative coverage and reimbursement schemes for pharmaceuticals. *J Manag Care Pharm* 5(8): 683-687.

Carnahan, S.J. (2007). Medicare's coverage with study participation policy: clinical trials or tribulations? *Yale J Health Policy Law Ethics* 7(2): 229-272.

Carino, T. and J. Gaffney (2012). Guidance 2.0 for coverage with evidence development: Striking the right chord. *Health Affairs* (blog), January 9.

<http://healthaffairs.org/blog/2012/01/09/guidance-2-0-for-coverage-with-evidence-development-striking-the-right-chord/>.

Carino, T., R.D. Williams, A.M. Colbert, and P. Bridger (2006). Medicare's coverage of colorectal cancer drugs: a case study in evidence development and policy. *Health Affairs* 25(5): 1231-1239.

Carino, T., S. Sheingold, and S.R. Tunis (2004). Using clinical trials as a condition of coverage. Lessons from the National Emphysema Treatment Trial. *Clinical Trials* 1: 1-14.

Chalkidou, C., A. Hoy, and P. Littlejohns (2007). Making a decision to wait for more research: When the National Institute for Health and Clinical Excellence recommends a technology only in the context of research. *J R Soc Med* 100(10): 453-460.

Claxton, K., S. Palmer, L. Longworth et al. (2012). Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development. *Health Technology Assessment* 16(46): 1-323.

Center for Medical Technology Policy (CMTP) (2011). *Coverage with Evidence Development: Key Issues*. Baltimore, MD: Center for Medical Technology Policy.

Center for Medical Technology Policy (CMTP) (2010). *Coverage with Evidence Development: A Conceptual Framework*. Baltimore, MD: Center for Medical Technology Policy.

College voor zorgverzekeringen (CVZ) (2012). *Conditional Reimbursement of Health Care*. Rapport Series No. 2012090234, April 6. Diemen, Netherlands: College voor zorgverzekeringen.

Dahm, L.L. (2008). *The Elusive “Reasonable and Necessary”: Making a Determination of Coverage with Evidence Development*.
[http://www.law.uh.edu/healthlaw/perspectives/\(LD\)CMSCoverageWithEvidence.pdf](http://www.law.uh.edu/healthlaw/perspectives/(LD)CMSCoverageWithEvidence.pdf).

Dhalla, I.A., S. Garner, K. Chalkidou, and P. Littlejohns (2009). Perspectives on the National Institute for Health and Clinical Excellence’s recommendations to use health technologies only in research. *International Journal of Technology Assessment in Health Care* 25: 272-280.

Eckermann, S. and A.R. Willan (2008). The option value of delay in health technology assessment. *Medical Decision Making* 28: 300-305.

Goeree, R., K. Chandra, J.E. Tarride, et al. (2010). Conditionally funded field evaluation: PATHs coverage with evidence development program for Ontario. *Value in Health* 13(Supp 1): S8-S11.

Hutton, J., P. Trueman, and C. Henshall (2007). Coverage with evidence development: An examination of conceptual and policy issues. *International Journal of Technology Assessment in Health Care* 24(4): 425-435.

Jaroslowski, S. and M. Toumi (2011). Market access agreements for pharmaceuticals in Europe: diversity of approaches and underlying concepts. *BMC Health Services Research* 11: 259.

Klemp, M., F.B. Fronsdal, and K. Facey (2011). What principles should govern the use of managed entry schemes? *International Journal of Technology Assessment in Health Care* 27: 77-83.

Levin, L., R. Goeree, M. Levine, et al. (2011). Coverage with evidence development: The Ontario experience. *International Journal of Technology Assessment in Health Care* 27: 159-68.

Levin, L., R. Goeree, N. Sikich, B. Jorgensen, M.C. Brouwers, T. Easty, and C. Zahn (2007). Establishing a comprehensive continuum from an evidentiary base to policy

development for health technologies: The Ontario experience. *International Journal of Technology Assessment in Health Care* 23(3): 299–309.

Lexchin, J. (2011). Coverage with evidence development for pharmaceuticals: a policy in evolution? *International Journal of Health Services* 41(2): 337-354.

Lindsay, M.J., B.A. Siegel, S.R. Tunis, B.E. Hillner, A.F. Shields, B.P. Carey, and R.E. Coleman (2007). The national oncologic PET registry: Expanded Medicare coverage for PET under coverage with evidence development. *American Journal of Roentgenology* 188(4): 1109-1113.

Longworth, L., J. Youn, L. Bojke, et al. (2013). When does NICE recommend the use of health technologies within a programme of evidence development?: A systematic review of NICE guidance. *Pharmacoeconomics* 31(2): 137-149.

McCabe, C.J., T. Stafinski, R. Edlin, and D. Menon (2010). Access with evidence development schemes: A framework for description and evaluation. *Pharmacoeconomics* 28(2): 143-152.

McCabe, C., K. Claxton, and A. O'Hagan (2008). Why licensing authorities need to consider the net value of new drugs in assigning review priorities: Addressing the tension between licensing and reimbursement. *International Journal Technology Assessment in Health Care* 24(2): 140-145.

Menon D, McCabe CJ, Stafinski T, Edlin R. (2010). Principles of design of Access with Evidence Development approaches. A consensus statement from the Banff Summit. *Pharmacoeconomics*, 28(2): 109-11.

Miller, F.G. and S.D. Pearson (2011). Linking insurance coverage for innovative invasive procedures with participation in clinical research. *Journal of the American Medical Association* 306(18): 2024-2025.

Miller, F.G. and S.D. Pearson (2008). Coverage with evidence development: Ethical issues and policy implications. *Medical Care* 46: 746-751.

- Mohr, P.E. and S.R. Tunis (2010). Access with Evidence Development: The US experience. *Pharmacoeconomics* 28(2): 153-162.
- Niezen, M., A. de Bont, E. Stolk, A. Eyck, L. Niessen, and H. Stoevelaar (2007). Conditional reimbursement within the Dutch drug policy. *Health Policy* 84: 39-50.
- Pietzsch, J.B. and M.E. Pate-Gornell (2008). Early technology assessment of new medical devices. *International Journal of Technology Assessment in Health Care* 24: 36-44.
- Redberg, R.F. (2007). Evidence, appropriateness, and technology assessment in cardiology: A case study of computed tomography. *Health Affairs* 26(1): 86-95.
- Retel, V.P., J.M. Bueno de Mesquita, J.M. Hummel, et al. (2009). Constructive technology assessment (CTA) as a tool in coverage with evidence development: The case of the 70-gene prognosis signature for breast cancer diagnostics. *International Journal of Technology Assessment in Health Care* 25(1): 73-83.
- Relyea-Chew, A. (2011). Ethical considerations with CMS's coverage with evidence development. *Journal of American College Radiology* 8(12): 838-841.
- Rutten, F., C. Uyl-de Groot, and A. Vulto (2009). Innovative payment systems for medicines in Europe. *EJHP Practice* 15: 60-62.
- Stafinski, T., C.J. McCabe, and D. Menon (2010). Funding the unfundable. Mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems. *Pharmacoeconomics* 28(2): 113-142.
- Towse, A. and L.P. Garrison (2010). Can't get no satisfaction? Will pay for performance help?: toward an economic framework for understanding performance-based risk-sharing agreements for innovative medical products. *Pharmacoeconomics* 28(2): 93-102.

- Trueman, P., D.L. Grainger, and K.E. Downs (2010). Coverage with evidence development: applications and issues. *International Journal of Technology Assessment in Health Care*, 26: 79-85.
- Tunis, S.R. and S. Pearson (2006). Coverage options for promising technologies: Medicare's "Coverage with Evidence Development". *Health Affairs* 25(5): 1218–1230.
- Tunis, S.R., T.V. Carino, and R.D. Williams (2007). Federal initiatives to support rapid learning about new technologies. *Health Affairs* 26: 140-149.
- Tunis, S.R. and K. Chalkidou (2007). Coverage with evidence development: A very good beginning, but much to be done. Commentary to Hutton et al. *International Journal of Technology Assessment in Health Care* 23(4): 432-435.
- Tunis, S.R. and D.M. Whicher (2009). The National Oncologic PET Registry: Lessons learned for coverage with evidence development. *J Am Coll Radiology* 6(5): 360-365.
- Tunis, S.R., R.A. Berenson, S.E. Phurrough, and P.E. Mohr (2011). Improving the Quality and Efficiency of the Medicare Program through Coverage Policy. Washington, DC: Urban Institute.
- Turner, J.A., W. Hollingworth, B. Comstock, and R.A. Deyo (2010). Comparative effectiveness research and policy: Experiences conducting a coverage with evidence development study of a therapeutic device. *Medical Care* 48(6 Suppl): S129-S136.
- Walker, S., M. Sculpher, K. Claxton, and S. Palmer (2012). Coverage with evidence development, only in research, risk sharing, or patient access scheme? A framework for coverage decisions. *Value in Health* 15(3): 570-579.
- Whicher, D.M., K. Chalkidou, I.A. Dhalla, L. Levin, and S. Tunis (2009). Comparative effectiveness research in Ontario, Canada: Producing relevant and

timely information for health care decision makers. *Milbank Quarterly* 87(3): 585-606.

Appendix B: *CED study interview guides*

Payer/HTA Body Questionnaire

Name of interviewee:

Title and affiliation of interviewee:

Date of interview:

Interviewer:

USE OF CED SCHEMES IN YOUR JURISDICTION

We are interested in better understanding the development and operations of CED programs and policies in different jurisdictions. The first several questions aim to obtain an informed overview of CED in your country.

1. When was the CED policy established?
2. What was the impetus for developing the policy?
3. Who was involved in its development and oversight?
4. What other stakeholders, if any, were involved in designing and implementing the policy?
5. Could you please describe the CED policy (e.g., criteria for CED; process for identifying candidate technologies; availability of different CED approaches, such as “only in research” or “only with research”; preferred study design/data collection method(s) or those used mostly frequently to date)? ***[Probe for answers to these domains if the interviewee does not address independently].***
6. In your opinion, is the CED process relatively standardized or is it guided more on a case-by-case basis?
7. Which types of technologies have the majority of CED policies been applied to date – drugs, devices, procedures, or a combination?
8. Can you give some examples of technologies that have undergone or are currently undergoing CED? ***[If interviewee is unable to recall example technologies, ask if they can point you to someone who might be able to address or to relevant source material, such as websites, publications, etc.]***
9. Prior to instigating a CED policy, to your knowledge is there some assessment of whether the benefits of additional research/evidence exceed the potential costs? ***[If the interviewee responds “yes”, request further information on the process. If the interviewee responds “no”, ask whether he or she believes this should be standard practice as part of overall policy.]***
10. What methodological approaches (e.g., RCTs, registries, etc.) are most frequently used in CED studies?
11. In your opinion, what are the advantages and disadvantages of different study design approaches?
12. Are their formal standards in place to establish when to modify or withdrawal coverage based on CED study findings?

13. In your opinion, what have been the greatest challenges in designing and implementing CED to date?
14. What changes are needed to ensure CED is more effective?
15. Are there proposals or plans in your country to modify current CED policies? If so, could you please provide a few examples?

CASES OF CED POLICIES APPLIED TO MEDICAL DEVICES

The following section is focused on examples where CED policies have been specifically applied to medical devices. To our knowledge, the following medical devices have undergone or are undergoing CED in particular countries: Spinal Cord Stimulation (SCS), Stenting and Aggressive Medical Management, Drug-Eluting Stents (DES), Implantable Cardioverter Defibrillators (ICDs), Laparoscopic Surgery, and Transcatheter Aortic Value Replacement (TAVR/TAVI).

16. To your knowledge, what devices have undergone CED, either in your country or in others? *[If the interviewee is uncertain, ask if they can point you to someone who might be able to address or to relevant source material, such as websites, publications, etc.]*

We would now like to ask some specific questions about one or two of these medical devices to obtain a better understanding of CED studies applied to medical devices. Would you be able to respond to such questions? [If interviewee responds “yes”, proceed. If interview responds “no”, ask if they can point you to someone who might be able to address or to relevant source material, such as websites, publications, etc.]

17. What was the impetus behind the CED policy and when was it established?
18. What were the terms of coverage?
19. Who was involved in the CED study?
20. How was the study funded?
21. What study approach was used?
22. Approximately how many patients were enrolled/involved?
23. What were the main endpoints collected in the study?
24. What were the primary outcomes of the study? In other words, did the study sufficiently address the uncertainty in the initial evidence base and overall conclusions about the value of the technology? Did the new evidence lead to adoption, modified adoption, or withdrawal of the device?

25. In your opinion, what were the main challenges (technical, procedural, political) to the use of CED in this particular case? Do you feel that any of these challenges are unique to the device? [*If interviewee responds “yes”, ask he or she to expound.*]
26. What would have improved the effectiveness or impact of this particular case of CED?

POTENTIAL MEDICAL DEVICES TO UNDERGO FUTURE CED

Given the fairly limited number of devices to have undergone CED to date, we would like to turn the discussion to devices that may be good candidates for CED.

27. In your opinion, what devices would be viable candidates for CED and why?
28. What CED approach would be beneficial to address any areas of uncertainty about the adoption and use of this (these) device (devices)?
29. Are there any aspects of this (these) particular medical device (devices) that would require special consideration either in determining whether it (they) should undergo CED or in designing the CED study?

GENERAL QUESTIONS ON THE APPLICATION OF CED TO MEDICAL DEVICES

We would like to close with a few general questions on the application of CED to medical devices.

30. What CED methodological approaches are most appropriate for medical devices? [*Probe to get a sense of whether the diversity of devices would necessitate different CED approaches.*]
31. In your opinion, what are the key methodological challenges to CED in the context of medical devices?
32. In your opinion, do medical devices introduce different challenges to CED compared to other types of technologies, such as drugs? [*If interviewee responds “yes”, ask he or she to expound and follow-up by asking whether there should be distinct CED frameworks/processes for different technology types*].
33. In general, what implications do you believe CED has or could have on innovation?
34. Is there anything else you would like to mention about the use of CED policies for medical devices?

Thank you very much for participating in our interview. We greatly appreciate your time and input.

[Ask interviewee if they would be interested in receiving our final study findings and note response (circle): Yes/No]

Industry Representative Questionnaire

Name of interviewee:

Title and affiliation of interviewee:

Date of interview:

Interviewer:

CASES OF CED POLICIES APPLIED TO MEDICAL DEVICES

We would first like to talk about examples where CED policies have been applied to medical devices. To our knowledge, the following medical devices have undergone or are undergoing CED in particular countries: Spinal Cord Stimulation (SCS), Stenting and Aggressive Medical Management, Drug-Eluting Stents (DES), Implantable Cardioverter Defibrillators (ICDs), Laparoscopic Surgery, and Transcatheter Aortic Value Replacement (TAVR/TAVI).

1. To your knowledge, what devices have undergone CED, either in your country or in others? *[If interviewee is uncertain, ask if they can point you to someone who might be able to address or to relevant source material, such as websites, publications, etc.]*

We would now like to ask some specific questions about one or two of these medical devices to obtain a better understanding of CED studies applied to medical devices. Would you be able to respond to such questions? [If interviewee responds “yes”, proceed. If interview responds “no”, ask if they can point you to someone who might be able to address or to relevant source material, such as websites, publications, etc.]

2. What was the impetus behind the CED policy and when was it established?
3. What were the terms of coverage?
4. Who was involved in the CED study?
5. How was the study funded?
6. What study approach was used?
7. Approximately how many patients were involved/enrolled?
8. What were the main endpoints collected in the study?
9. What were the primary outcomes of the study? In other words, did the study sufficiently address the uncertainty in the initial evidence base and overall conclusions about the value of the technology? Did the new evidence lead to adoption, modified adoption, or withdrawal of the device?
10. In your opinion, what were the main challenges (technical, procedural, political) to the use of CED in this particular case? Do you feel that any of these challenges are unique to the device? *[If interviewee responds “yes”, ask he or she to expound.]*
11. What would have improved the effectiveness or impact of this particular case of CED?

POTENTIAL MEDICAL DEVICES TO UNDERGO FUTURE CED

Given the fairly limited number of devices to have undergone CED to date, we would like to turn the discussion to devices that may be good candidates for CED.

12. In your opinion, what devices would be viable candidates for CED and why?
13. What CED approach would be beneficial to address any areas of uncertainty about the adoption and use of this (these) device (devices)?
14. Are there any aspects of this (these) particular medical device (devices) that would require special consideration either in determining whether it (they) should undergo CED or in designing the CED study?

GENERAL QUESTIONS ON THE APPLICATION OF CED TO MEDICAL DEVICES

We would like to close with a few general questions on the application of CED to medical devices.

15. What CED methodological approaches are most appropriate for medical devices? [*Probe to get a sense of whether the diversity of devices would necessitate different CED approaches.*]
16. In your opinion, what are the key methodological challenges to CED in the context of medical devices?
17. In your opinion, do medical devices introduce different challenges to CED compared to other types of technologies, such as drugs? [*If interviewee responds “yes”, ask he or she to expound and follow-up by asking whether there should be distinct CED frameworks/processes for different technology types.*]
18. In general, what implications do you believe CED has or could have on innovation?
19. Is there anything else you would like to mention about the use of CED policies for medical devices?

Thank you very much for participating in our interview. We greatly appreciate your time and input.

[Ask interviewee if they would be interested in receiving our final study findings and note response (circle): Yes/No]

Academics/Policy Analysts Questionnaire

Name of interviewee:

Title and affiliation of interviewee:

Date of interview:

Interviewer:

USE OF CED SCHEMES IN YOUR JURISDICTION

We are interested in better understanding the development and operations of CED programs and policies in different jurisdictions. The first several questions aim to obtain an informed overview of CED in your country.

1. When was the CED policy established?
2. What was the impetus for developing the policy?
3. Who was involved in its development and oversight?
4. What other stakeholders, if any, were involved in designing and implementing the policy?
5. Could you please describe the CED policy (e.g., criteria for CED; process for identifying candidate technologies; availability of different CED approaches, such as “only in research” or “only with research”; preferred study design/data collection method(s) or those used mostly frequently to date)? ***[Probe for answers to these domains if the interviewee does not address independently].***
6. In your opinion, is the CED process relatively standardized or is it guided more on a case-by-case basis?
7. Which types of technologies have the majority of CED policies been applied to date – drugs, devices, procedures, or a combination?
8. Can you give some examples of technologies that have undergone or are currently undergoing CED? ***[If interviewee is unable to recall example technologies, ask if they can point you to someone who might be able to address or to relevant source material, such as websites, publications, etc.]***
9. Prior to instigating a CED policy, to your knowledge is there some assessment of whether the benefits of additional research/evidence exceed the potential costs? ***[If the interviewee responds “yes”, request further information on the process. If the interviewee responds “no”, ask whether he or she believes this should be standard practice as part of overall policy.]***
10. What methodological approaches (e.g., RCTs, registries, etc.) are most frequently used in CED studies?
11. What are the advantages and disadvantages of different study design approaches?
12. Are their formal standards in place to establish when to modify or withdrawal coverage based on CED study findings?

13. In your opinion, what have been the greatest challenges in designing and implementing CED to date?
14. What changes are needed to ensure CED is more effective?
15. Are there proposals or plans in your country to modify current CED policies? If so, could you please provide a few examples?

CASES OF CED POLICIES APPLIED TO MEDICAL DEVICES

The following section is focused on examples where CED policies have been specifically applied to medical devices. To our knowledge, the following medical devices have undergone or are undergoing CED in particular countries: Spinal Cord Stimulation (SCS), Stenting and Aggressive Medical Management, Drug-Eluting Stents (DES), Implantable Cardioverter Defibrillators (ICDs), Laparoscopic Surgery, and Transcatheter Aortic Value Replacement (TAVR/TAVI).

16. To your knowledge, what devices have undergone CED, either in your country or in others? *[If interviewee is uncertain, ask if they can point you to someone who might be able to address or to relevant source material, such as websites, publications, etc.]*

We would now like to ask some specific questions about one or two of these medical devices to obtain a better understanding of CED studies applied to medical devices. Would you be able to respond to such questions? [If interviewee responds “yes”, proceed. If interview responds “no”, ask if they can point you to someone who might be able to address or to relevant source material, such as websites, publications, etc.]

17. What was the impetus behind the CED policy and when was it established?
18. What were the terms of coverage?
19. Who was involved in the CED study?
20. How was the study funded?
21. What study approach was used?
22. Approximately how many patients were involved/enrolled?
23. What were the main endpoints collected in the study?
24. What were the primary outcomes of the study? In other words, did the study sufficiently address the uncertainty in the initial evidence base and overall conclusions about the value of the technology? Did the new evidence lead to adoption, modified adoption, or withdrawal of the device?

25. In your opinion, what were the main challenges (technical, procedural, political) to the use of CED in this particular case? Do you feel that any of these challenges are unique to the device? [*If interviewee responds “yes”, ask he or she to expound.*]
26. In your opinion, what would have improved the effectiveness or impact of this particular case of CED?

POTENTIAL MEDICAL DEVICES TO UNDERGO FUTURE CED

Given the fairly limited number of devices to have undergone CED to date, we would like to turn the discussion to devices that may be good candidates for CED.

27. In your opinion, what devices would be viable candidates for CED and why?
28. What CED approach would be beneficial to address any areas of uncertainty about the adoption and use of this (these) device (devices)?
29. Are there any aspects of this (these) particular medical device (devices) that would require special consideration either in determining whether it (they) should undergo CED or in designing the CED study?

GENERAL QUESTIONS ON THE APPLICATION OF CED TO MEDICAL DEVICES

We would like to close with a few general questions on the application of CED to medical devices.

30. What CED methodological approaches are most appropriate for medical devices? [*Probe to get a sense of whether the diversity of devices would necessitate different CED approaches.*]
31. In your opinion, what are the key methodological challenges to CED in the context of medical devices?
32. In your opinion, do medical devices introduce different challenges to CED compared to other types of technologies, such as drugs? [*If interviewee responds “yes”, ask he or she to expound and follow-up by asking whether there should be distinct CED frameworks/processes for different technology types.*]
33. In general, what implications do you believe CED has or could have on innovation?

34. Is there anything else you would like to mention about the use of CED policies for medical devices?

Thank you very much for participating in our interview. We greatly appreciate your time and input.

[Ask interviewee if they would be interested in receiving our final study findings and note response (circle): Yes/No]

Appendix C: Details of the interview sample

Type of Expert	Number Interviewed	% of Total Sample
Payer/HTA body	7	32%
Industry representative	5	23%
Policy analyst/academic	10	45%
TOTAL	22	100%

Country Represented	Number Interviewed	% of Total Sample
Canada	1	5%
France	1	5%
Germany	2	10%
Netherlands	3	15%
Switzerland	2	10%
United Kingdom	4	20%
United States	4	20%
Other*	5	25%
TOTAL	22	100%

*Informant associated with an international, mainly European, representation.

Appendix D: Select CED schemes in Canada, UK, and the US

CANADA			
Disease Area	Product	Scheme	Impact on Coverage Policy
Multiple cancers	PET for head and neck cancers	Questions remained about the clinical utility of pre-surgery following radiation therapy. The Ontario Clinical Oncology Groups conducted a prospective cohort study to address area of uncertainty.	Study findings suggest no clinical utility and, consequently, PET was not insured for head and neck cancer indications.
Vascular disease	Endovascular Abdominal Aortic Aneurysm Repair (EVAR)	Potential safety issues resulting from endoleaks and uncertain cost-effectiveness. To address these issues, PATH, with an Ontario Academic Health Services Centre, conducted a prospective, observational study.	Study confirmed no issues regarding potential for endoleaks and found that EVAR only cost-effective for high surgical risk patients. EVAR funded for high, but not low, risk patients.
UNITED KINGDOM			
Disease Area	Product	Scheme	Impact on Coverage Policy
Cervical cancer	Liquid-Based Cytology (LBC)	Insufficient evidence to justify nationwide introduction of LBC. Coverage limited to patients participating in large-scale pilot implementation studies carried out to evaluate the effectiveness, costs, and practical implications of the introduction of LBC technology into the cervical screening programme.	Treatment was eventually recommended as the main way of preparing samples of cervical cells for cervical screening.
Lymphoma (follicular non-Hodgkin's)	Rituximab	Technology was only recommended for last-line therapy in the context of the construction of a case series study of past and new patients, in order to determine with more certainty its effectiveness in this indication.	Rituximab later recommended in combination with select agents as an option for the treatment of symptomatic stage III and IV follicular lymphoma in previously untreated people.
Hip disease	Metal-on-Metal (MoM) Hip Resurfacing Arthroplasty	Where hip resurfacing arthroplasty is considered appropriate, NICE recommended that the procedure should only be performed in the context of ongoing data collection on both the clinical- and cost-effectiveness of the technology, and ideally via a UK national joint registry. Until long-term evidence is available, NICE recommended that surgeons choose a device for MoM resurfacing for which there is at least 3 years' evidence.	Use of MoM followed via UK National Joint Registry. Analysis of the registry (through 2012) demonstrated that the total revision rate for the MITCHTRH System total hip replacements is higher than acceptable by NICE. Restrictions were later placed on the use of MITCH TRH acetabular cups/MITCH TRH modular heads in combination with

			uncemented femoral stems.
Breast cancer	Taxanes	Use of taxanes for adjuvant treatment of early breast cancer should be limited to patients enrolled in RCTs.	Docetaxel was eventually recommended for breast cancer, while paclitaxel was not (Chalkidou, 2006).
Colorectal cancer	Oxaliplatin and Irinotecan	Neither treatment was recommended for first-line treatment for advanced colorectal cancer except as a part of a clinical trial (Chalkidou, 2006).	Both treatments were eventually recommended for wide spread use (Chalkidou, 2006).
Chronic Pain of Ischaemic Origin	Spinal Cord Stimulation	Only recommended in the context of research designed to generate robust evidence (preferably RCTs) about the benefits of spinal cord stimulation (including pain relief, functional outcomes, and quality of life) compared with standard of care.	No studies to date.
UNITED STATES			
Disease Area	Product	Scheme	Impact on Coverage Policy
Cognitive Impairment	FDG-Positron Emission Tomography (PET) Scan	An FDG-PET scan covered in patients with mild cognitive impairment or early dementia in the context of a clinical trial.	Study began in 2006 and is still in the process of recruiting patients. Study estimated to be completed in early 2016.
Emphysema	Lung Volume Reduction Surgery (LVRS)	Rapid growth in surgery volume despite little evidence supporting its safety and effectiveness prompted CMS to suspend payments and co-sponsor a nationwide RCT to evaluate the procedure. LVRS covered only as part of a clinical trial (NETT) (Carino et al., 2004c).	The main outcome of the study was mixed, resulting in questions about the meaningfulness of the results to substantiate meaningful benefit (Ramsey and Sullivan, 2005). The economic analyses also suggested that the procedure had relatively poor cost-effectiveness in the short run. Nevertheless, CMS agreed to cover LVRS.
Hearing Loss	Cochlear Implant	CMS may cover cochlear implantation for treatment of hearing loss in the context of an approved clinical trial. Patients must have hearing test scores of greater than 40% and less than or equal to 60% and only when the provider is participating in, and patients are enrolled in, either an FDA-approved investigational device trial or a prospective, controlled, comparative trial approved by CMS.	To date, no studies have been approved.

Tachyarrhythmia's	Implantable Cardioverter Defibrillators (ICDs)	Published trials led to broad consensus regarding the clinical benefits of ICDs in appropriately selected patients. However, there was agreement that important questions remained about the benefits and risks in specific patient subgroups (Tunis and Pearson, 2006). In particular, existing trial data could not identify the majority of patients for whom ICDs were unlikely to fire. Therefore, CMS issued a CED policy to expand coverage of ICDs to a requirement to submit data to a national registry.	Registry established that the risk of in-hospital complication rates was lower for ICD implantation performed by electrophysiologist than for other physician specialty types. However, a different registry design was needed to address remaining evidence gaps. A new registry was designed to address these issues, which is still in progress.
Artificial Hearts	Heart Disease	Artificial heart devices covered only when provided to Medicare beneficiaries when the device is implanted as part of a CMS approved clinical study.	Studies still ongoing.
Chronic low back pain (CLBP)	Transcutaneous Electrical Nerve Stimulation (TENS)	Coverage allowed for TENS only when the patient is enrolled in an approved RCT, using validated and reliable instruments, within three years (prior to June 2015). Studies must be designed so that patients in the control and comparison groups receive the same concurrent treatments and either sham (placebo) TENS or active TENS intervention.	Currently, no clinical trials have been approved by CMS.

Source: Authors' compilation based on literature review (mainly from the organisations' websites).