This report is part of the RAND Corporation research report series. RAND reports present research findings and objective analysis that address the challenges facing the public and private sectors. All RAND reports undergo rigorous peer review to ensure high standards for research quality and objectivity.
Regulatory cultures and research governance

Catriona Manville, Petal Jean Hackett, Salil Gunashekar, Molly Morgan Jones

Prepared for the English Department of Health
The research described in this document was prepared for the English Department of Health.

RAND Europe is an independent, not-for-profit research organisation whose mission is to improve policy and decision making for the public good. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

© Copyright 2013 RAND Corporation

Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Copies may not be duplicated for commercial purposes. Unauthorized posting of RAND documents to a non-RAND website is prohibited. RAND documents are protected under copyright law. For information on reprint and linking permissions, please visit the RAND permissions page (http://www.rand.org/publications/permissions.html).

Published 2013 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
Westbrook Centre, Milton Road, Cambridge CB4 1YG, United Kingdom
RAND URL: http://www.rand.org
RAND Europe URL: http://www.rand.org/rand_europe
To order RAND documents or to obtain additional information, contact Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org
Preface

This report is intended to help improve understanding of health research governance in the UK by exploring the regulatory practices and cultures in other countries and sectors. It is intended to explore what can be learned from a comparative study of the practice of those who are subject to regulatory requirements in the health research, medical drugs, environmental and financial sectors. The report is informed by a review of a small subset of literature which is particularly relevant to this question, and focuses on different elements of regulation and regulatory governance for each of the different sectors.

The intended audience for this report is policymakers at the English Department of Health, regulators within the Health Research Authority, researchers, research institutions, local NHS Trusts, sponsors of research and members of the public with an interest in research governance.

RAND Europe is an independent not-for-profit policy research organisation that aims to improve policy and decision making in the public interest, through research and analysis. RAND Europe’s clients include European governments, institutions, non-governmental organisations and firms with a need for rigorous, independent, multidisciplinary analysis. This report has been peer-reviewed in accordance with RAND’s quality assurance standards.

For more information about RAND Europe or this document, please contact:

Molly Morgan Jones
RAND Europe
Westbrook Centre
Milton Road
Cambridge CB4 1YG
United Kingdom
Tel: +44 (0)1223 353329
Email: mmjones@rand.org
Contents

Preface ................................................................. iii
Table of figures ....................................................... ix
Table of tables ......................................................... xi
Summary ................................................................. xiii
Acknowledgements ................................................... xvii

CHAPTER 1  Introduction ......................................................... 1
  1.1  Background ......................................................... 1
  1.2  Methodology ....................................................... 2
    1.2.1  Task 1: Scoping and identification of comparative frameworks ...... 2
    1.2.2  Task 2: Literature search ........................................ 3
    1.2.3  Task 3: Literature review ......................................... 3
    1.2.4  Task 4: Comparative analysis .................................... 3
    1.2.5  Task 5: Reporting .................................................. 3
  1.3  Comparative framework ........................................... 3
  1.4  Structure of the report .............................................. 4

CHAPTER 2  The regulation of research within the English healthcare system ..................... 5
  2.1  Introduction ........................................................ 5
  2.2  Approval to commence research ................................... 5
  2.3  Academy of Medical Sciences report .................................. 7
  2.4  Introduction of the Health Research Authority ......................... 7

CHAPTER 3  International health research governance systems ........................................... 9
  3.1  Overview of research systems and regulation .............................. 9
    3.1.1  Australia ......................................................... 10
    3.1.2  Brazil ............................................................ 12
    3.1.3  Canada .......................................................... 13
    3.1.4  China .............................................................. 15
    3.1.5  India ............................................................... 16
    3.1.6  Russia ............................................................. 17
    3.1.7  The USA ......................................................... 18
    3.1.8  Summary ......................................................... 20
3.2 The challenge of multi-site trials .............................................................. 21
3.3 Case study 1: Australia ........................................................................... 21
3.4 Case study 2: Canada .............................................................................. 23
3.5 Case study 3: The USA .......................................................................... 25
3.6 Conclusion .............................................................................................. 28

CHAPTER 4 \hspace{10pt} Comparative framework 1: medical drug regulation .......... 31
4.1 Introduction ............................................................................................ 32
4.2 Key regulatory events in US medical drug approval ................................. 32
\hspace{10pt} 4.2.1 Specific event: Thalidomide ..................................................... 33
\hspace{10pt} 4.2.2 Specific event: the AIDS crisis ............................................... 35
4.3 Key regulatory events in UK medical drug approval ................................. 36
\hspace{10pt} 4.3.1 Specific event: Thalidomide crisis ........................................... 36
\hspace{10pt} 4.3.2 Specific event: European directives and the European Medicines
\hspace{10pt} Agency ......................................................................................... 36
4.4 Regulatory attitudes to pharmacovigilance .............................................. 38
4.5 Lessons from medical drug regulation .................................................... 39

CHAPTER 5 \hspace{10pt} Comparative framework 2: Environmental regulation .......... 41
5.1 The rise of environmental governance .................................................... 42
5.2 Environmental governance and contaminated land management in the
\hspace{10pt} UK ........................................................................................................ 44
5.3 Environmental governance and proactive corporate environmental
\hspace{10pt} management .................................................................................. 46
5.4 Environmental governance and stakeholder engagement in the
\hspace{10pt} consideration of environmental risk decision-making ...................... 47
5.5 Lessons from environmental governance ............................................ 49

CHAPTER 6 \hspace{10pt} Comparative framework 3: financial services sector .......... 51
6.1 Introduction ............................................................................................ 51
6.2 Recent regulatory events in the US financial services sector ...................... 52
\hspace{10pt} 6.2.1 Specific event: the Savings and Loan crisis ................................ 53
\hspace{10pt} 6.2.2 Specific event: the 2007 global financial crisis ......................... 54
6.3 Recent regulatory events in the UK financial services sector ...................... 56
\hspace{10pt} 6.3.1 Specific event: establishment of the Financial Services Authority .... 57
\hspace{10pt} 6.3.2 Specific event: The 2007 global financial crisis ......................... 58
6.4 Lessons from financial services .............................................................. 60

CHAPTER 7 \hspace{10pt} Comparative analysis: lessons for research governance in the
\hspace{10pt} UK ................................................................................................. 63
7.1 Educational initiatives ............................................................................ 65
7.2 Transparency among actors .................................................................... 64
7.3 Trust and confidence in the system and system actors............................ 65
Table of figures

Figure 6-1: The post-FIRREA US financial regulatory system.......................... 54
Figure 6-2: The post-Dodd-Frank US regulatory system................................ 56
Figure 6-3: The tripartite arrangement for UK financial regulation............... 57
Figure 6-4: UK financial regulatory system, post-FSA restructuring............. 60
Table of tables

Table 3-1: Summary of research governance models by country................................. 20
Health research governance in the UK has been the subject of much debate and discussion, particularly over the past few years, as existing regulatory approaches and processes have been consolidated and reconsidered. This has been done in response to a growing concern about overly bureaucratic and duplicative review processes for the approval of research in the UK health system, including the approval of clinical trials. This raises questions about why the regulated sector, including the applicants, sponsors, research institutions and local NHS Trusts, are responding to the regulation in ways that lead to duplication. This research was done in order to improve our understanding of the impact of regulation on health research governance, specifically in relation to the behaviour of the regulated sector. When framed in this way, the question can be extended beyond the health research sector and we can look comparatively at the behaviour of regulated sectors in different areas and countries.

Therefore, this study is centred on the key research question:

What can be learned from a comparative study of the practice of those who are subject to regulatory requirements across sectors and countries?

This study is informed by a review of a small subset of the literature around regulation and regulatory governance in four sectors: health research, medical drug approval, environmental risk regulation and financial services. Once the review of each sector was completed, a comparative analysis drew out lessons which might be applied in the context of health research governance in the UK. Thus, this review is not intended, neither does it claim, to be an exhaustive study of regulation and related literature within each of the sectors. Rather, the sectors and literature within them were selected in the interest of identifying areas which would lend themselves to comparative analysis in the context of health research governance.

This review of health research systems in Australia, Brazil, Canada, China, India, Russia and the USA found that countries use different models to regulate and review research, although most systems have dual components of both decentralised (local) and centralised (national) processes. A further focus on the responses and changes introduced in Australia, Canada and the USA to cope with multi-site trials found that the following mechanisms were used or suggested as ways to affect the way that the regulatory system is received and responded to by stakeholders:

- use accreditation systems to instil trust into review boards which receive decisions from other review boards
• provide certification of staff to provide the required mutual trust in others’
decision making – this can be particularly effective when introduced through a
national training programme
• encourage reciprocal agreements to accept others’ decisions
• increase the transparency of the decision making process to build trust between
parties, building on shared approval systems or making use of standard operating
procedures
• increase interaction and communication between committees to establish
relationships and trust between the individuals involved
• provide education and encourage the use of evidence to understand the relative
risk compared to the hypothetical
• evaluate or audit the current system to determine overall success – thus producing
more confidence in the system.

This review of the medical drug industry in the UK and the USA focuses on selected key
events over the postwar period. Specifically, we consider the Thalidomide and Acquired
Immune Deficiency Syndrome (AIDS) crises, the introduction of European directives and
the establishment of the European Medicines Agency. These events provide good natural
experiments for understanding the effect of policy changes on the regulated sector, as the
immediate changes in the regulatory stance provide a benchmark against which changes in
behaviour can be examined. Based on this review, the authors suggest that the ways in
which changes to the regulatory system for medical drugs are received and responded to by
stakeholders offer lessons for the regulation of health research. These include the following.

• Managing seemingly inconsistent regulatory objectives carefully so that they are
complementary and not contradictory. Failing to properly balance objectives may
result in a cycle of increasing and relaxing standards, eroding confidence in the
regulator. Further, if the guidelines resulting from this balancing exercise are not
fed through a regulatory body to the relevant executing units, they may experience
delays in conducting reviews.
• Developing clear, consistent guidelines on what constitutes a poor outcome (such
as unnecessary risk to human lives) and how it will be handled, including setting
out, where relevant, criminal and civil liabilities for non-compliance.
• Leveraging public pressure to enhance industry compliance with regulations.
• Considering regulatory actions in the context of the overall landscape in which
firms operate: for example, return on investment in drug development is affected
heavily by developments in the intellectual property rights landscape. Regulators
must consider the likely impact of new regulatory actions on firms’ ability to
extract profit under intellectual property rights, in order to understand and/or
predict compliance behaviours better.
The review of the environmental regulation literature showed that there has been a broader move towards a more dispersed model of environmental governance, as opposed to top-down regulation. This has meant that a range of different mechanisms are used to ensure that environmental risks are minimised, including the following.

- Harness the role of public trust and confidence. Demand from the public for environmental accountability was not only an early driver of regulatory action, but also is a current driver of proactive corporate environmental management.
- Equally, harness consumer demand, particularly where the government has less direct control over the behaviour of the regulated sector.
- Ensure there that is no misalignment of regulatory and actor philosophies, which can pose a threat to implementation of regulation and hence present challenges for effective and efficient governance responses.
- Use education, training and capacity-building to encourage actors to engage with each other, and to foster understanding of the views of different stakeholders in order to create a system that is seen by all as more legitimate and effective.
- Use incentives as the actors become more dispersed and behaviours more difficult to control. Here it is necessary to take into account the motivations of different actors and shape incentives accordingly.

Finally, this analysis of regulatory developments in the financial services industry in the UK and the USA includes the Savings and Loan crisis in the USA, the formation and dissolution of the Financial Services Authority in the UK and the global financial crisis. It examines how immediate changes in the regulatory stance have affected the behaviour of the regulated sector. It is suggested that the lessons to be learned include the following.

- Focus less on the form of the regulator and more on consistent application of the rules, along with closure of regulatory gaps. Much has been written about the pros and cons of functional regulation in the USA versus the umbrella regulator in the UK, but both have failed to treat the contributing factors to the global crisis.
- Pay careful attention to potential asymmetry in regulatory effects. Increasing the burden on one segment versus another (such as banks relative to non-banks, in the US case) or one location versus another, might lead to inter-segment competition and makes the entire system more fragile.
- Pace changes to regulation and allow adequate time for adjustment. The true impact of changes in regulation are observed fully only with a lag, and this effect is greater, the more complex the regulation. Comprehensive changes should be given longer to mature than simple changes.
- Recognise the trade-off between targeted and comprehensive regulations. Complex regulations (for example, the Financial Services and Markets Act 2000 in the UK and the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 in the USA) may produce the required shifts in firm compliance and risk.
behaviour, but also may result in overlapping and potentially contradictory rules, making it easier for firms to undertake risky behaviour.

- Carefully consider whether contradictory policy objectives can be executed by a single body or whether balancing among the objectives should be carried out at a higher level.

Finally, this cross-cutting comparative analysis identified five common elements or initiatives across the sectors which could be considered further, in relation to health research governance in the UK:

- increased provision of educational initiatives to improve awareness and training among actors
- transparency and promotion of a culture of openness between researchers, sponsors, trusts, institutions and the public as to the decision making and approval processes which are followed
- together with education and transparency, development of additional mechanisms to foster trust within the system. Formal (accreditation schemes or memoranda of understanding) and informal (networking, relationship-building) mechanisms should be considered
- consideration of the regulatory structure, including where trade-offs may need to be made to align regulatory philosophies and objectives
- use of incentives, in particular the role of the public in creating a demand for research, should be explored – this includes determination of different indicators and metrics that actors can be evaluated against, such as research publications, trials hosted or number of new participants recruited.
We would like to thank Arnav Kapur for very helpful initial research for this project which was conducted during his internship at RAND Europe. Dr Janet Wisely, Dr Frank Wells and Dr Hugh Davies at the Health Research Authority were gracious with their time in providing early insights into the project. We also wish to thank Sachi Yagyu, RAND librarian, for conducting the initial literature searches and Rosanna Jeffries, Research Assistant at RAND Europe, for her help with references and final formatting of the document. Finally, the authors wish to acknowledge Dr Saba Hinrichs and Dr David Kryl for their critical and constructive comments on earlier versions of this document during the quality assurance process.

This is an independent report commissioned and funded by the Policy Research Programme in the Department of Health. The views expressed are not necessarily those of the Department.
1.1 **Background**

Research governance in the UK has been the subject of much debate and discussion, particularly over the past few years, as existing regulatory approaches and processes have been consolidated and reconsidered. This has been done in response to a growing concern about overly bureaucratic and duplicative review processes for the approval of research in the UK health system, including the approval of clinical trials. The concern was perhaps most prominently articulated by the Academy of Medical Sciences in a 2010 report, published in 2011, detailing a series of recommendations to streamline the service. One key recommendation was the formation of a new Health Research Agency which would attempt to rationalise the regulation and governance of all health research. This national service would ensure that governance checks were completed only once, that reviews were conducted in a timely manner, and that general and specialist ethical approval were combined into one system (Academy of Medical Sciences, 2011). It would work closely with other research regulatory bodies and organisations, in particular the Medicines and Healthcare Products Regulatory Agency (MHRA), the National Research Ethics Service (NRES) which would be subsumed into the Health Research Authority (HRA) and the National Institute for Health Research (NIHR).

Although the Health Research Agency was designed to complete these functions and has been in place since late 2011, there is still a perception that the system is not as efficient as it could be. For example, it is suggested that duplication in ethical approval processes could be reducing the speed with which decisions are made.¹ This raises questions about how and why the regulated sector, including applicants, sponsors and research institutions, is responding to the regulation: in other words, what is the impact of regulation on research governance, specifically in relation to the behaviour of the regulated sector? When framed in this way, the question can be extended beyond the health research sector and we can look comparatively at the behaviour of regulated sectors in different areas and countries.

Therefore, this study is centred on the key research question:

What can be learned from a comparative study of the practice of those who are subject to regulatory requirements across sectors and countries?

¹ This emerged from a conversation with the Department of Health in the scoping phase for this project.
Informed by a review of the broader literature around regulation and regulatory governance, this report will look at the behaviour of regulated sectors in different areas and countries, and at what might be learned in the context of health research governance in the UK – particularly regarding how to influence the behaviour of the regulated sector and its internal governance within, and in spite of, the regulation that exists. It is worth noting that in all the sectors the primary focus of this study is on regulation of products, broadly interpreted and inclusive of health research, that are being released into a wider market, where regulation affects the behaviour of actors involved in production. This is complementary to, but distinct from, regulation of the way that services are delivered: for example, regulation of healthcare itself. Finally, for the purposes of this study, we use definitions provided by the Department of Health (DH) and take ‘regulation’ to mean both the legislation and the action of regulators to enforce it, whereas ‘governance’ refers to what goes on in terms of internal processes within the regulated sector, to ensure their own compliance with the law.

1.2 Methodology

Before discussing the methodological approach, it is worth noting briefly that this approach was used in favour of others which would have involved a more intensive primary data collection effort. Specifically, there are a few alternative methodologies which were considered, but ultimately not pursued.

First, short, semi-structured interviews with different National Health Service (NHS) trusts could be used to explore the ‘how’ and ‘why’ of their behaviours and responses to the regulation of research, including their perceptions and lack of trust in the centralised aspects of the research governance system. This approach was rejected on the basis that others are having such conversations with NHS Trusts and a desire to reduce the burden on trusts in taking time away from their daily work.

Second, it was considered whether an empirical study into the factors that led people to have confidence in the system would be worthwhile. Such a review might focus on specific ‘mistakes’ that were made in the past, or highlight instances where dual review was thought to pick up on an element that was missed by someone else. However, this was rejected on the basis of value for money and the consideration that this may not add value to the study.

Finally, due to the desire to highlight comparative analysis, it was thought that any sort of empirical study which adequately captured all sectors would take time and be costly. Therefore, value for money is highest, with a conceptual literature review that focused on secondary analysis of a selection of other studies of regulation and governance which would usefully inform the research questions. There were five key stages to the research, each of which is now described in turn.

1.2.1 Task 1: Scoping and identification of comparative frameworks

This task identified the main areas of comparison for the review, and selected three main sectors and international regulatory frameworks for comparison: the financial services sector, medical drug regulation and environmental risk regulation.
1.2.2 **Task 2: Literature search**

This task included developing keywords and search terms for the literature review. Literature searches were carried out in Scopus, Pubmed and Web of Knowledge, as well as grey literature sources, in order to identify the most relevant literature for each framework. The following search strings were used in the published literature, while in the grey literature modified search strings were used (due to the simpler nature of most web-based data sources):

- regulation AND (Comparative OR behavior OR governance OR risk) AND (financial services OR environmental OR health and safety)
- research regulation AND (clinical trials OR (Institutional Review Boards (IRBs) AND USA))
- (“clinical trial” OR “clinical trials”) AND (“ethical review” OR “ethical reviews”) AND approv*
- “research governance” AND (health or financ* OR environment*)
  research AND governance AND regulatory AND (health OR financ* OR environment*)

1.2.3 **Task 3: Literature review**

After the literature search, a detailed full text review of each paper was carried out to identify relevant data and information on each comparative framework, including insights into cultural influences, challenges in implementation, effects on the system (for example, lack of growth) and lessons learned.

1.2.4 **Task 4: Comparative analysis**

The comparative analysis looked across all frameworks and involved workshops between the researchers to brainstorm comparative findings for the study.

1.2.5 **Task 5: Reporting**

The final report summarises the main findings, in addition to providing detailed summaries of each comparative framework researched.

1.3 **Comparative framework**

A comparative framework was developed to analyse the behaviour of regulated sectors in the selected different areas and countries, and to draw out lessons for the context of research governance in the UK. This framework will draw on sectors within the health system which may provide useful in-sector comparisons, as well as sectors outside it, in order to provide useful cross-sector comparisons. In selecting the sectors and countries for comparison, the following criteria were used as a guide:

- the presence of a comparative element, such as a different sector or country
- evidence of a recent change in structure or regulatory framework in order to provide insight into how behaviours within the sector change in response to regulation, and where there are opportunities to affect behaviour change early on
- similar regulatory mechanisms to health research to ensure that the potential for confounding variables (such as self-governance) is ruled out
• regulation that focused on the ‘products’ within a sector, as opposed to regulation of how services such as healthcare are delivered.

Based on this set of criteria, initial scoping of the literature and internal expert advice, the following sectors or areas were selected for comparison:

• research governance in other health systems
• medical drug regulation and the pharmaceutical sector
• regulation of environmental risk
• the financial services sector and responses to regulatory changes.

1.4 **Structure of the report**

The remainder of this report is set out as follows. Chapter 2 discusses the research regulatory system in the UK. Chapter 3 describes the systems of research governance in other countries, focusing on Australia, Brazil, Canada, China, India, Russia and the USA. It focuses on the Australian, Canadian and American systems in particular, and outlines the mechanisms used by each to affect the way the regulatory system is received and responded to by stakeholders. Chapters 4, 5 and 6 explore changes in regulatory systems and responses to them by different groups of stakeholders within the medical drug, environmental risk and financial service sectors. Chapter 7 pulls together the analyses from the previous chapters, presents the findings of the cross-cutting comparative analysis across all sectors based on the literature reviewed, and identifies common themes which could be explored further for their applicability to research governance.
2.1 Introduction

Research is a key component of the UK healthcare system. It provides patients with early access to new treatments and innovations, and improves the quality and efficiency of the health service, thus bringing added benefits to the patient. Research must be regulated to ensure the safety of patients and to facilitate research. In 2005, The DH published a Research Governance Framework for Health and Social Care which outlines the principles of governance that apply to research conducted on humans (Department of Health, 2005). The dignity, rights, safety and well-being of participants is listed in the Research Governance Framework as the primary consideration for any research study. Therefore, it is essential that due scientific and ethical consideration be applied to all studies, and that approval is received before the study commences. Ethical approval is conducted by research ethics committees (RECs) in the UK, which provide an impartial and independent opinion on whether the proposed study complies with recognised ethical standards. Researchers are then required to keep the REC informed of the study’s progress.

2.2 Approval to commence research

Obtaining permission to conduct a research study in the UK can require approval from a variety of bodies to ensure the validity and viability of the trial and the safety of participants. In England,\(^2\) since its formation in 2006 the National Institute for Health Research (NIHR) has created the infrastructure and facilities to ensure that the NHS is a valuable and viable partner for research (Academy of Medical Sciences, 2011). The UK Clinical Research Network (UKCRN) is a part of NIHR and is tasked with helping to improve researchers to deliver clinical studies through:

- introducing effective systems to reduce the time taken to receive NHS permission for the conduct of commercial and non-commercial clinical research

\(^2\) The majority of the processes discussed in this chapter are specific to the English system. Where they are UK-wide, we state this. However, a comprehensive review of all four systems within the UK was beyond the scope of this research.
• creating an infrastructure to allow researchers access to the facilities and support that they require to undertake the study
• directing researchers toward research capacity and local patient populations to ensure that participant targets are achieved (National Institute for Health Research, 2012).

Clinical research networks have been established in the four geographical regions of the UK: England, Scotland, Wales and Northern Ireland. The NIHR Clinical Research Network Coordinating Centre (NIHR CRN CC) is based in both Leeds and London and supports the networks.

The process of study approval is centralised through the NIHR Coordinated System for gaining NHS Permission (CSP), which supports the application and approvals process for NIHR clinical research network portfolio studies. This system is managed by UKCRN. Within this portfolio, trials have access to NHS support for research and training, and receive an international standard randomised control trial number (ISRCTN). Since 2008, the information required by the different review bodies is captured centrally through the Integrated Research Application System (IRAS), streamlining the process for permissions and approvals and thereby reducing duplication.

Generally, there are two categories of approval required for any given study: research and development (R&D) permission, including site-specific assessment, (SSI), and ethics approval. Ethical approval is required from the National Research Ethics Service (NRES), which operates a common UK-wide system for the ethical review of health and social care research.

Currently there are 80 RECs in England. RECs are made up of clinical members from the fields of medicine, nursing and science, lay members and specialists: for example, pharmacists and radiation advisers. They consider three topics: the validity of the research; the welfare of the research participants; and the dignity of the research participants.

R&D approval is required from the NHS organisations hosting the study for any research undertaken within the NHS. This includes the involvement of NHS facilities, NHS patients, their tissue, data or samples. The required information can be broken down into a study-wide form and site-specific information forms for each site.

Additional approvals may be required depending on the type of study and the materials used. For example, if embryos are used, the Human Fertilisation and Embryology Authority (HFEA) must approve the study: this is granted through a research licence for up to three years, and decisions can take up to three months. This too is coordinated through IRAS. Other review bodies hosted through IRAS include:

• Administration of Radioactive Substances Advisory Committee (ARSAC)
• Gene Therapy Advisory Committee (GTAC)
• Medicines and Healthcare products Regulatory Agency (MHRA)
• Ministry of Justice
• NHS/Health and Social Care (HSC) R&D offices
• NRES/NHS/HSC RECs
• National Information Governance Board (NIGB)
• National Offender Management Service (NOMS)
• Social Care Research Ethics Committee (Integrated Research Application System, 2012).

If the research study is a clinical trial, a clinical trials application and permission from the MHRA, the competent authority in the UK under the EU Clinical Trials Directive, is required.

The order of applying for approvals depends on the type of study. For example, single-centre studies complete the NHS Trust R&D submission prior to the MHRA and ethics submission, whereas multi-centre trials obtain MHRA and ethics approval prior to approval from the NHS Trust offices involved.

2.3 **Academy of Medical Sciences report**

In January 2011, the Academy of Medical Sciences (AMS) reported its findings from a year’s comprehensive review of research within the NHS, commissioned by the government, involving human participants and their tissue or data. The key findings of the report, *A New Pathway for the Regulation and Governance of Health Research* (Academy of Medical Sciences, 2011) were as follows:

• research is stifled by a complex and bureaucratic regulatory environment
• regulations should safeguard patients and facilitate research
• streamlining of the regulatory and governance pathway is required
• a cultural change is required in the NHS to embed health research as a core function.

On these grounds, the report made a series of 17 recommendations including the introduction of an independent body to rationalise the regulation and governance of all health research. It was hoped also that the changes would lead to a cultural change which would ensure that research is valued within the NHS, and that there are improvements to the clinical trials environment in the UK.

2.4 **Introduction of the Health Research Authority**

One key recommendation from the AMS report was the formation of a new Health Research Agency, which would attempt to rationalise the regulation and governance of all health research. This national service would provide a single point of access and contact for researchers during the approvals process, and ensure that research governance checks were completed only once, reviews were conducted in a timely manner, and general and specialist ethical approval were combined into one system (Academy of Medical Sciences, 2011). The agency would work closely with the other research regulatory bodies and
organisations, in particular the MHRA, NRES (which would be subsumed into the Agency as a core service) and NIHR.

In response to the report, the government produced a series of commitments in its *Plan for Growth* (Department for Business Innovation and Skills, 2011) and in December 2011, the Health Research Authority (HRA) was set up. One of the initiatives in *Plan for Growth* is that from 2013, NIHR funding to providers of NHS services will become conditional on a 70-day benchmark to recruit first patients for trials. This means that all research approval at the NHS Trust level must be conducted in less than 70 days. A meeting held by the AMS, Cancer Research UK and The Wellcome Trust in May 2012 discussed current successes and issues with the new system, including concerns about how to improve and streamline research governance. Generally it was agreed that the rapid establishment of the HRA was well received. However, there was concern that NHS Trusts were still duplicating scientific or ethical review that only needed to happen once, and that there may be other repetition of approval throughout the system. It is thought that this is leading to delays in the system.

Through comparative analysis of other health systems and other sectors, this report aims to complement the current initiatives of the DH around improving research governance, and to provide areas for further exploration which may help increase the efficiency and effectiveness of the system.
CHAPTER 3  International health research governance systems

Chapter summary
Countries use different models to regulate and review research. This chapter aims to describe the different regulations and governance mechanisms used within seven countries: Australia, Brazil, Canada, China, India, Russia and the USA. Most systems have dual components of both decentralised (local) and centralised (national) processes. In particular, this chapter focuses on comparing the responses that regulatory systems have employed to cope with the rise of multi-site trials: this looks at the situation in Australia, Canada and the USA. The aims of this change are varied across each country, but generally have been done to reduce duplication, inefficiency and delays in the trial initiation and recruitment. These changes are intended to make the process smoother for researchers, sponsors and review committees.

The common mechanisms used to affect the way the regulatory system is received and responded to by stakeholders, include the following:

- use accreditation systems to instil trust in review boards which receive decisions from other review boards
- provide certification of staff to provide the required mutual trust in others’ decision making – this can be particularly effective when introduced through a national training programme
- encourage reciprocal agreements to accept others’ decisions
- increase transparency of the decision making process to build trust between parties, building on shared approval systems or making use of standard operating procedures
- increase interaction and communication between committees to establish relationships and trust between the individuals involved
- provide education and encourage the use of evidence to understand the relative risk compared to the hypothetical
- evaluate or audit of the current system to determine overall success, thus producing more confidence in the system.

3.1 Overview of research systems and regulation
Research regulations and governance models vary from country to country, but are widely considered to be essential elements of modern clinical medicine in order to protect human subjects from undue harm.
There are three main types of review for research involving human subjects (centralised, decentralised and dual), with six basic models within them, although the models are not mutually exclusive (Eckstein, 2005).

1. Devolved review – a REC within an institution conducts a review of the ethical components of the research that is to be conducted within that institution.
2. Mutual acceptance between committees or institutions – REC reviews are accepted reciprocally by other RECs or other institutions conducting research.
3. Shared review – multiple institutions create a single REC for the review of research within their region.
4. Delegation or agency model – one institution’s REC acts on behalf of another institution through an agreement or contract.
5. Centralised review – a central REC conducts ethical approval for all research. This can be based on location (region, state or country) or specialisation.
6. Ad hoc review – institutions or RECs form a contract with another to undertake review on a case-by-case basis.

The focus in the list above is on RECs because these are the most consistently utilised form of review in each country. It is worth noting that this is only one classification scheme, and others have proposed different ways of thinking about the nature of research governance in order to understand the behaviours within the system more fully. For example, Fitzgerald and Phillips (2006) summarise the basic models as centralised (one form, one review), decentralised (centralised and local committees) or dual (everything is considered locally), but then delineate further between administrative and ethical reviews. This latter distinction separates out the substantive ethical and technical review of research, from the administrative tasks involved in review, such as designing and providing application forms, management of submitted forms and other clerical activities. This, they argue, is a useful way of understanding the different discourses which highlight the perceived barriers and challenges to centralised and decentralised systems. This is something that will be returned to at times in the subsequent analysis, as this study seeks to understand how and why actors within the regulatory system develop different governance responses.

However, as the next section will show, other types of review are present, and the nature of additional review processes is contingent upon the system and country. It also will be examining if countries differentiate between site-specific, ethical and scientific review. To illustrate these review processes, the regulations and governance mechanisms within seven countries – Australia, Brazil, Canada, China, India, Russia and the USA – have been briefly analysed.3

3.1.1 Australia

In Australia, review of research is governed by the ‘National Statement on ethical conduct in human research’ (Australian Government, 2007). Human RECs (HRECs) conduct the scientific and ethical review of research. They are overseen by the National Health and

---

3 Access to information, particularly in non-English speaking countries, was mixed and so the comprehensiveness of the information varies slightly by country.
Medical Research Council (NHMRC). Independent review of clinical research became mandatory in 1985, although many committees (previously called Institutional Ethics Committees; IECs) were in existence prior to this date. HRECs are established within institutions that conduct research, most commonly universities and hospitals. Currently there are more than 200 HRECs registered with the Australian Health Ethics Committee, with 37 percent of those in the public health sector, 18 percent in the private health sector and 24 percent in the university sector (Australian Government National Health and Medical Research Council, 2004).

HRECs are made up of at least eight members. As far as possible there should be an equal number of men and women, and at least one-third of the members should be from outside the institution for which the HREC is reviewing research. The committee should be made up of the following:

- a chairperson
- at least two laypeople (one of each sex) who have no affiliation with the institutions and do not currently engage in medical, scientific, legal or academic work
- at least one person with knowledge and current experience of professional care, counselling or treatment of people (such as a nurse)
- at least one person with a community pastoral care role (such as a minister of religion)
- at least one lawyer
- at least two people with current research experience that is relevant to the research proposals to be considered.

These final members can be selected as required from an established pool of members with expertise.

The primary role of the HREC is to protect the welfare and rights of participants in research through ensuring that studies involving human participants are ethically acceptable and in accordance with relevant standards and guidelines.

The Australian system has evolved into its current form from a more decentralised system where review was conducted at individual institutions. Where research was proposed to take place at multiple institutions, researchers and committees were not required to work together, but they were strongly encouraged to do so. The 1999 National Statement (subsequently updated in 2007) provides guidelines on conducting multi-site trials and encouraged researchers to:

- come to a prior agreement with one HREC that it would take the primary role in the ethical and scientific assessment of a protocol
- inform each HREC of all the other Australian sites at which the research is being proposed or conducted.

---

4 Section 5.1.30. of the National Statement on Ethical Conduct in Human Research.
It also encouraged HRECs to:

- avoid unnecessary duplication of ethical review by seeking to ascertain if the same protocol has been reviewed elsewhere
- communicate with and give advice to or receive advice from any other HREC
- accept scientific or technical assessment by another organisation
- review and adopt the reasons for ethical approval or disapproval of another HREC in reaching its own decision.

However, this encouragement-based system had problems and did not have much uptake among the research review community. The issues are summarised in this quote from a paper provided for a 2005 inter-jurisdictional forum to discuss challenges to the system:

The changes also took insufficient account of the previous seven years of IEC concern about legal exposure in clinical trial review and of the increased independence of institutions. By 1999, three decades of a robust tradition of independence, exacerbated by seven years of strengthened institutional autonomy in clinical trial review, left HRECs in major medical research institutions deeply resistant to devolving or sharing any of their responsibilities for ethical review. (Breen, 2005, p. 11)

In response to these concerns, in October 2006, the Australian Health Ministers’ Advisory Council (AHMAC) agreed to implement the national system described above, which would facilitate the recognition of a single scientific and ethical review process within and across all Australian jurisdictions. In response to this and in order to address potential overlaps in the system, the Harmonisation of Multi-centre Ethical Review (HoMER) project was funded by the national government in 2007 (National Health and Medical Research Council (NHMRC, 2012).

### 3.1.2 Brazil

Clinical research trials in Brazil are regulated by the National Research Council under Resolutions No 196/96 and 251/97 (Rodrigues and Kesselring, 2008). These regulations focus primarily on setting out the ethical and logistical aspects of research studies. For a clinical trial to be approved, it must go through several stages of approval which are coordinated by the Ministry of Health, but require separate sets of approval from local and central ethics bodies and from the National Health Surveillance Agency (ANVISA), a regulatory body equivalent to the US Food and Drug Administration or the MHRA in the UK. ANVISA's role in the approval process is to focus on protocol methodology and monitoring of the clinical trial data as it is available, as well as to give approval for the importation and licensing of the drugs themselves, once the trial is completed.

The National Commission for Research Ethics (Comissão Nacional de Ética em Pesquisa; CONEP) in Brazil was established in 1996. CONEP is a constituent body within the National Health Council (CNS) in the Ministry of Health. CONEP is responsible for accrediting and coordinating RECs (Comitês de Ética em Pesquisa, CEPs) at the institutional level (that is, CEPs are established in universities, medical institutions, etc. and are equivalent to institutional review boards (IRBs). Each CEP comprises a panel of members with varied expertise (for example, health scientists, social scientists, jurists, bioethicists, etc.). Rodrigues and Kesselring (2008) reported that there were 539 CEPs in Brazil, the majority (86 percent) of which are concentrated in São Paolo and Rio de
Janeiro. This number had gone up to 598 by 2010 (Lopes, 2011). Lopes lists the specific aims of the CONEP/CEP system in great detail, but in general, the system is responsible for reviewing and monitoring different ethical matters related to any kind of scientific research involving human participants.

The overarching management and governance structure of the health research system in Brazil is led by the Ministry of Health; however, the Ministry of Science and Technology and the Ministry of Education also have coordinating functions within the wider research system. Individual states have their own ministries of health and science and technology, which implement regulations at a local level and oversee the actions of local CEPs (Alger et al., 2009).

In summary, for a clinical research trial to be approved, it must follow these steps.

1. Translation of the protocol and associated documents into Portuguese.
2. Review and approval by the coordinating site’s CEP (local level).
3. Review and approval by CONEP after the coordinating site CEP has given its approval.
4. Review and approval by each individual CEP. Although this is done in parallel with CONEP’s review, local CEPs usually wait until CONEP has approved the study before giving their own approval (Rodrigues and Kesselring, 2008).
5. Separate approval from CONEP is required also if the trial is funded by a foreign sponsor.
6. Review and approval by ANVISA after all the CEPs and CONEP have given their approval.
7. Final approval for the study to commence is issued.

It is interesting to note that until very recently, the ethics approval system in Brazil involved hard copies of documents being posted to the relevant entities. To overcome the exceedingly large time delays associated with this process, an online platform called the Registro Brasileiro de Ensaios Clínicos (ReBEC) was launched in November 2011 by the Minister of Health, which is meant to facilitate researchers in submitting documents as well as regularly tracking progress during the project review phase (Andrews and Pizolato, undated). It is anticipated that with this new system there will be a significant reduction in the duration of the ethics approval process. Other amendments to the regulatory process have been under discussion in recent years, including improving the capabilities of local CEPs. Under this new system, coordinating CEPs for a study will renew the licence of other sites only if their CEPs have adequate facilities, regulatory knowledge and qualified evaluators. The aim of this new system is to strengthen the CEPs to the point that CONEP will no longer have to issue its own separate approval. In addition, ANVISA has moved to reviewing trial protocols alongside CONEP and local CEPs, in an attempt to reduce delays in the system (Rodrigues and Kesselring, 2008).

### 3.1.3 Canada

Ethical review is a requirement of federal funding in Canada. Depending on the funding source, there are different requirements for review. The top three federal funding agencies (Canadian Institutes of Health Research, CIHR; Natural Sciences and Engineering Research of Canada, NSERC; and Social Sciences and Humanities Research Council of Canada, SSHRC) formed the interagency advisory panel of research ethics in 2001.
Through this they developed the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS). Therefore, research ethics boards (REBs) at institutions that receive funds from one or more federal research agencies must apply the TCPS. Other organisations have now adopted the TCPS, for example Health Canada.

In Canada, the review of research is conducted by REBs. These are multidisciplinary committees established by institutions to undertake the ethical review of research projects affiliated with that institution. REBs have financial and administrative independence from their institutions. The TCPS requires these REBs to consist of at least five members, including:

- both men and women, where two members have broad expertise in the methods or area of research covered by the REB, one member who is knowledgeable in ethics
- one member from the community that the REB serves.

In addition, REBs are advised to have:

- one member knowledgeable in the law, to alert the REB to legal issues and their implications (this is compulsory for those reviewing biomedical research)
- ad hoc members, who can be added if extra expertise is required.

In addition to ethical review, REBs conduct scholarly review, which is the evaluation of the academic or scientific merits of the research.

The review board considers the risk of harm and the benefits of the research. The TCPS requires that foreseeable harm should not outweigh the anticipated benefits; the research participants must not be subjected to unnecessary risks of harm; and the benefits associated with the research must be made as great as possible and the risks minimised. The TCPS requires REBs to adopt a proportionate approach to ethics review. The level of scrutiny required for the research review is related to the risk carried. There are several levels of ethics review:

- full REB review involves a face-to-face meeting
- expedited REB review involves an individual or subgroup of the REB departmental review is used for the ethical review of undergraduate projects carried out within formal course requirements.

The TCPS mandate is a living document. In the original mandate, individual REBs were ‘responsible for the ethical acceptability of research undertaken within its institution’, even when the study was a multi-site study. An additional chapter in the updated 2010 Second Edition of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCP2); Government of Canada, undated) details multiple models for ethics review in the case of multi-site studies. These options aim to provide flexibility, as the application depends on the context of the research and the ethical norms and practices of the relevant research discipline. These models are:

---

5 For further information, see Government of Canada (undated) Panel on Research Ethics.
• independent ethics review by several REBs
• research ethics review delegated to an external specialised or multi-institutional REB
• reciprocal REB review.

The independent review option can occur concurrently or sequentially within each REB. Central REBs can be used under option 2, which focus on particular types of research. One example, the Ontario Cancer Research Ethics Board (OCREB), is a centralised, oncology-specific REB concentrating specifically on the governance surrounding cancer clinical trials (Chaddah, 2008). This enables the ethics board to find and retain appropriately qualified board members for the specific subject. Further to approval from the central REB, individual submissions to the trial sites were approved in parallel. Reciprocal REBs occur when multiple institutions agree to accept each other’s REB’s review results. An example of this is the province-wide reciprocal approval initiative in Alberta where, in February 2011, the six REBs signed a reciprocity agreement (Alberta Health Services, 2012).

3.1.4 China

In China there are numerous government bodies that are responsible for the management (and funding) of biomedical research. For example, the State Food and Drug Administration (SFDA) oversees the approval of all clinical trials for pharmaceutical drug products, while the National Expert Ethics Committee, whose operations are overseen by the Chinese Ministry of Health, is responsible for ethics review.

The official report of the China–UK Research Ethics (CURE) Committee published in 2009 contains a detailed analysis of the research ethics guidelines prevalent in China (Medical Research Council, 2009). For the purpose of the study, ‘research ethics’ included both research governance and bioethics matters. The CURE report additionally examined the implementation of the research ethics guidelines in specific research settings in China.

The overall responsibility for regulating ethics of research involving human participants lies with the Ministry of Health. The National Committee of Bioethics functions only in an advisory capacity to the ministry. There are numerous laws, regulations and guidelines governing biomedical research, and in particular, research ethics. One such regulation is the Regulation on Ethical Review of Biomedical Research Involving Human Subjects proposed by the Ministry of Health in 2007, which contains information on informed consent procedures and outlines the rules to be followed when setting up ethics committees.

In terms of organisational structure, China has IRBs similar to the US research system, which monitor and review research according to the national regulations. These review boards have to approve all biomedical research before it can be conducted. The administration of the IRBs is supervised at the state or provincial level by the

---

6 The CURE Committee was convened by the UK Medical Research Council in partnership with the Foreign and Commonwealth Office and the China National Centre for Biotechnology Development, with the primary aim of studying various issues related to research ethics that arose in the context of UK–China collaborative biomedical research involving human participants.
corresponding departments of health. Each IRB is expected to have at least five members (both men and women) and ‘members from non-medical profession, law and other institutions’ (Institutional Review Board Guidebook, 1993).

When dealing with clinical trials of new drug or medical products, the procedures involved in the application and approval become more complex. Within the decision making process, site approval and technical assessment of the trial are two separate review processes. The former is done at the provincial level, while the latter is conducted by a national body.

3.1.5 India

In India, the Department of Health Research (DoHR)\(^7\) is responsible for providing guidance on various issues related to research governance. These include the ethical issues encountered in medical and health research, and there is a commitment by the Government of India to encourage ethical practices in these areas. Indeed, ethics is highlighted as one the primary values in its recent plans to create an overarching National Health Research System (NHRS)(Department of Health Research, 2011).

The DoHR funds the Indian Council of Medical Research (ICMR), which is the primary organisation responsible for formulating, coordinating and promoting biomedical research in the country. The ICMR has issued ethical guidelines for biomedical research with respect to both ‘human experimentation’ (Indian Council of Medical Research, 2006) and ‘laboratory animal welfare’ (Indian Council of Medical Research, undated). The ICMR guidelines stipulate that in order to protect the rights of participants and safeguard their welfare, it is compulsory for all research proposals that involve human participants to be cleared by an appropriately constituted institutional ethics committee (IEC). IECs are equivalent to the IRBs in the US system. It is interesting to note that independent ethics committees (IEC(Ind)) exist to assist researchers who are not attached to a particular institution, or who are employed by institutions that do not have any ethics committees.

The three key responsibilities of an IEC are as follows:

1. to safeguard the rights, well-being and dignity of the prospective research participants
2. to ensure that universal ethical values and international scientific standards are expressed in terms of the local community values and customs
3. to support the development of a research community that is able to cater to local healthcare requirements.

Generally, in terms of its composition, an IEC is formed of eight to 12 members, of which a minimum of five are required to form a quorum for the purpose of decision making. To obtain different perspectives, IECs are multidisciplinary in composition, consisting of, for example, medical scientists, clinicians, legal experts, social scientists, philosophers and laypeople. There is no formal training provided to these members; however, they are encouraged to keep themselves well informed regarding developments in national and

---

\(^7\) The Department of Health Research (http://www.dhr.gov.in/) was created in 2007 as a subsidiary of the Ministry of Health and Family Welfare (http://www.mohfw.nic.in/), signalling the Government of India’s desire to ensure that health research plays an increasingly important role in framing health policy.
international ethical practices. They can do this by attending orientation courses on appropriate topics that are conducted either by their colleagues or by ‘constituted bodies’. In the specific context of drug trial reviews, the ICMR guidelines recommend that IEC members are trained in Good Clinical Practice.

Before a particular research project is initiated, the IEC has the responsibility of reviewing the research proposal and ensuring that the proposal has been scientifically evaluated. The IEC’s secretary screens the proposal for its completeness and, based on the risks involved, categorises it one of into three groups:

1. **exempt from review** – this is for proposals which pose ‘less than minimal risk’ to the research participants
2. **expedited review** – this is for proposals which pose ‘no more than minimal risk’ to the research participants
3. **full review** – this includes all proposals which pose more than minimal risk to the participants, as well as proposals which cannot be processed under the first two categories.

In addition to the initial reviews, IECs are responsible for continuously monitoring approved programmes of research to ensure that they comply with the stipulated ethical guidelines. These periodic reviews may occur at intervals of six to 12 months.

It is worth noting that it is a mandatory requirement for all clinical trials to comply with the ICMR ethical guidelines according to the Drugs and Cosmetic Act 2002 and the Medical Council of India Act 2002 (UNESCO, 2010). Finally, after much speculation and inordinate delays, the Biomedical Research Human Subjects Promotion and Regulation Bill, which has been drafted by the ICMR, is set to be introduced into parliament during its winter session in November 2012 (Shankar, 2012).

### 3.1.6 Russia

The Ministry of Health and Social Development in Russia is responsible for the establishment and/or accreditation of RECs (European Forum for Good Clinical Practice (EFCGP), 2011). The ministry also supervises the ethics committees and maintains quality standards. As a constituent of the Ministry of Health and Social Development, the Council of Ethics is the primary central committee. The Council of Ethics consists of 17 members, with a minimum of two-thirds constituting a quorum. Additionally, there are a number of local ethics committees formed, for example, in hospitals, universities and so forth.

In order to proceed with an ethical review of a clinical trial, the sponsor is required to submit a review request to the Council of Ethics. In addition to this, the principal investigator of the clinical trial is required to submit the review request to the appropriate local REC. Once a particular research programme has been approved, it is the responsibility of the local ethics committee to monitor the clinical trial: for a long-term clinical trial, it is expected that an intermediate report is submitted.

---

8 The ICMR guidelines define minimal risk as ‘one which may be anticipated as harm or discomfort no greater than that encountered in routine daily life activities of the general population or during performance of routine physical or psychological examinations or tests’.

9 There was less information available in English for this system.
The members of the Council of Ethics are required to attend an internal training course on research ethics. However, this is not a requirement for members of local ethics committees. Members can also enrol for relevant courses at universities.

3.1.7 **The USA**

The system of research governance in the USA is grounded in local, institutionally based review. While estimates vary, there are likely to be anywhere from 4,000 to 6,000 local IRBs in the USA at any given time (Wood, Grady and Emanuel, 2002). These review boards provide the task of reviewing and approving all research that is going to be conducted at or by researchers at the institution. IRB review is required for all research which is federally funded, or which may involve testing new medical products that will be ultimately submitted to the US Food and Drug Administration (FDA) for approval. Since the majority of biomedical research in the USA is federally funded (The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979), most institutions sign documents similar to memoranda of understanding that all of their research, regardless of funding, is subject to the same procedures of review.

The history of the regulatory system stems from the Belmont Report, issued in 1979 in response to many scandals in the use of human participants in research in an unethical and problematic manner (for example, the Tuskegee syphilis trial and the controversial use of pregnant woman to test the experimental drug Thalidomide) (The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979). While the Belmont Report set out the basic ethical principles which should guide research, another report issued at the same time set out the context for IRB review that exists today, ensuring that there is accountability at a local level and to assure local oversight of the research (Fleischman, 2005).

The design, purview and conduct of local IRBs are all set out in various federal regulations and codes, in particular Title 45, Part 46 of the Code of Federal Regulations of 2009 (US Department of Health & Human Services, 2009). These regulations set out the types of research which are subject to human participant research protection and the requirements for the registration, composition, conduct and review procedures of IRBs. Most relevant for this discussion are the composition, conduct and review procedures.

IRBs must be composed of at least five members with varying backgrounds and disciplinary expertise (US Department of Health & Human Services, 2009). Moreover, in addition to the ability to perform adequate technical and scientific review, the IRBs must ‘be able to ascertain the acceptability of proposed research in terms of institutional commitments and regulations, applicable law and standards of professional conduct and practice’ (US Department of Health & Human Services, 2009). Furthermore, IRBs must include at least one lay member and at least one member who is not affiliated with the institution.

In relation to conduct and the nature of the review, IRBs are required by law to consider the following points:

- the risks to participants are minimised
- the risks to participants are reasonable in relation to the anticipated benefits
- the selection of participants is equitable
• informed consent procedures are appropriate
• data monitoring plans are in place, when appropriate
• additional safeguards are in place for vulnerable participants.

Thus, IRBs are effectively a decentralised mechanism of enforcing centralised regulations for the conduct of research, including scientific, ethical and usually site-specific considerations in the approval of research. Although not specifically stated in the regulations, it appears that local, site-specific assessment seems to be a part of the IRB process.\textsuperscript{10} The present study found references to site-specific assessments of capability and concerns in both discussions of IRB review and in reference to internal standard operating procedures within individual institutions which also might be considered as a part of, or in parallel with, the IRB review. IRBs review the scientific quality of the study; however, this assessment is also conducted in the funding body approval to support the study, as well as in the case of US Food and Drug Administration (FDA) approval, if a new medical product is being tested in a clinical trial.

There has been a recent trend in the USA towards centralised IRBs, (CIRBs) and more is said about this in the USA case study below. However, these only apply to certain types of trials in specific disease areas or with specific disease populations, and so are not a common feature of the US system.

\textsuperscript{10} This was the case for IRBs for which we were able to find and review the submission forms. It is worth noting, though, that without a more qualitative review involving interviewing those involved in the system, it was difficult to ascertain exactly how and at which point site-specific assessment comes into consideration.
### Summary

Table 3-1 summarises the main features of each system.

#### Table 3-1: Summary of research governance models by country

<table>
<thead>
<tr>
<th>Country</th>
<th>Review structure</th>
<th>Review models used</th>
<th>Ethical review?</th>
<th>Site-specific review?</th>
<th>Scientific review?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>Centralised</td>
<td>• Centralised review at a state and national level</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes, conducted by HREC</td>
</tr>
</tbody>
</table>
| Brazil | Dual | • Devolved review by coordinating CEP  
• Mutual acceptance of coordinating CEP  
• Centralised review by CONEP  
• Centralised review by ANVISA | Yes | Yes | Yes |
| Canada | Decentralised | • Devolved review by individual REBs  
• Central review by a specialist REB  
• Reciprocal review | Yes | Yes, conducted by the REB | Scientific validity reviewed prior to REB review |
| China | Dual | • Devolved review by individual IRBs, supervised at provincial level  
• Central review of clinical trials of new drugs | Yes | Yes | Yes |
| India | Decentralised | • Devolved review by individual IECs | Yes | Unknown | Yes, responsibility of IEC to check |
| Russia | Dual | • Central review by Council of Ethics  
• Decentralised oversight of research by local ethical review boards | Yes | Unknown | Unknown |
| USA | Dual | • Devolved review by individual IRBs  
• Centralised review by CIRBs in some cases  
• Centralised review for drug or medical approval | Yes | Yes | Yes, conducted by the IRB |
With this summary completed, this report now turns its attention to one of the driving forces of change in the regulation of research: the challenge of multi-site trials.

3.2 The challenge of multi-site trials

Over the past several years there has been a move towards trials occurring at multiple sites within a country or even internationally. However, without a centralised process there are issues with conducting multiple reviews on a single study. For example, applicants report variation in the information required by, and verdicts received from, individual boards when provided with the same information (Silverman, Hull and Sugarman, 2001). In addition, there is redundancy in the process, and it has been argued that the cost of multiple reviews does not produce added benefits.

Within the last decade many countries including Australia, Canada, the UK and the USA have introduced initiatives, although at different speeds and at different levels, which move some aspects of approval from being conducted at individual sites to more centralised systems. The aims of this change are varied across each country (see the individual country discussions below), but generally have been done to reduce duplication, inefficiency and delays in trial initiation and recruitment. These changes are intended to make the process smoother for researchers, sponsors and review committees.

However, alongside the calls for centralisation and streamlining are deep-seated tensions in places such as Canada and the USA, which have long had decentralised systems. Although researchers and sponsors would like to streamline the process, institutions and often the public and media fear that the loss of local oversight will mean that there is less control over the nature and conduct of the trial. Concerns about liability and responsibility also remain. These and other concerns are discussed in greater detail in the following sections for the cases of Australia, Canada and the USA.

3.3 Case study 1: Australia

In Australia there are two sets of challenges and opportunities to consider, given recent changes to the system and the introduction of a more centralised research regulatory system, as described above. Prior to the centralised system, it was recognised that research spanning more than one jurisdiction faced problems. The barriers within the systems included duplication of review, overall burden of work and different standards and skills between HRECs (Pittman, 2007). In addition, the issue of indemnity and insurance around reviewing clinical trials was raised several times in the literature.

In response to the issues in the current system, a forum was held in 2005 to identify opportunities for addressing these and to discuss the introduction of a centralised system by the National Health and Medical Research Council (NHMRC). One issue in particular which was addressed was that of liability. At the 2005 inter-jurisdiction forum it was noted that there are legal requirements obligating HRECs to undertake ‘in-house’ ethical review, as they are concerned for the legal liabilities that may arise if they rely on the review of another HREC. A forum paper states that ‘institutions should recognise the extremely small risk which attaches to HREC review and take a practical, rather than theoretical approach to this risk’ (Zeps, Frew and Kelly, 2005, p. 20). Over the past decade there have
been very few cases of researchers being sued in Australia, and no known cases of HREC members being sued. In order to increase trust that the information obtained from other HRECs or institutions is robust, a national accreditation system or independent quality assurance benchmarking system was suggested.

This centralised system was piloted and implemented initially in New South Wales in 2005 and 2007 respectively. The pilot allowed an initial scientific assessment by a central shared scientific assessment committee (SSAC) (Frew and Martlew, 2005); the final scheme allowed an individual HREC to be determined as the lead HREC for the study, and to give central scientific and ethical approval covering all public health sites in New South Wales. Following central approval, each site submits a site-specific assessment (SSA) of its capability to conduct the trial, covering the suitability of the research, required facilities, staff expertise and site-specific regulatory considerations. It was noted during the pilot that despite centralisation of the overall process, there was still a need for the site appraisal to be conducted at the institution. As with R&D offices at the NHS Trusts in the UK, some institutions in Australia have research offices which conduct the required site assessment. To standardise this process, Victoria and New South Wales have developed a site-specific information form: this template should be appropriate to all institutions, regardless of their research management systems.

This reform was received with mixed success by researchers using the process. For example, a case study in New South Wales comparing the length of time taken to conduct approval for a multi-site trial through the centralised process, with time taken to apply to the individual HRECs, showed that the median time for initial approval was less for the centralised trial than the non-centralised review system (89 days compared to 100 days respectively; Vajdic et al., 2012). However, the addition of site-specific assessment added an average of 60 days to the whole process, and therefore overall on average, approval of the trial through the centralised system took longer to complete. Despite this, the authors of the study did acknowledge that the centralised system had reduced duplication and the administrative burden of conducting an ethical review, as well as promoting consistency of site documentation.

The review of another state pilot (conducted in Victoria), assessing the acceptance of multi-centre trial approval, listed HRECs’ initial propensity to ‘forget’ the agreed text and processes to which they had signed up previously (Alt, 2005). This meant that duplication was still seen in the system, as HRECs continued to conduct their own review in addition to the lead committee review. However, as the pilot progressed and trust between the different committees was built up, all HRECs ceased this practice.

Following these pilots, in 2007 a report to the NHMRC discussed the steps required to create a national system of HRECs providing once-only ethical and scientific evaluation for research spanning multiple states and territories (Pittman, 2007). This system has now been implemented, but it has taken time to take hold and there are still challenges within it. Various initiatives described below have been used (both by regulators and those within the system) since the new structure was introduced, in an attempt to overcome some of these barriers of implementation.

Accreditation has been used to establish credibility for lead HRECs. In order to become a lead HREC in New South Wales, HRECs must be nominated for accreditation by the
chief executive of the New South Wales Public Health Organisation, and meet accreditation standards set by the director-general of New South Wales Ministry of Health. This accreditation model has been discussed on the national level, as well as the involvement of peer review and objective analysis. These less formalised elements of peer review and objective analysis allow greater flexibility in the system, and are deemed to encourage continuous improvement through education and training. The proposal recommends that lead HRECs are reviewed by suitably qualified and experienced people biannually. The criteria by which HRECs are reviewed are:

- membership of the HREC
- constitution of the HREC
- frequency of meeting
- workload
- conformity of operations with the HREC’s own standard operating procedure (SOP) as well as nationally provided SOPs.

In addition, there is an annual survey of researchers and sponsors that have used the HREC within the last year, with an emphasis on how positive the interaction was and whether the HREC added value to the research proposal under review. It is worth noting that the SOPs were suggested as a method to introduce harmonisation within the system. These should be available and followed by all participating HRECs.

There is a national ethics application form (NEAF) which is completed for all applications for ethical and scientific review. In 2012 the New South Wales Ministry of Health, and Queensland and Victoria Departments of Health, have signed a memorandum of understanding to introduce the mutual acceptance of ethical and scientific review of multi-centre clinical trials undertaken in public health organisations across the three states. This memorandum of understanding is an important step in establishing a reciprocal system and helping to build trust between sites, so that duplication is minimised.

3.4 Case study 2: Canada

In Canada, an essentially decentralised review system, there are many issues associated with undertaking multiple site review at the institutional level (Interagency Advisory Panel and Secretariat on Research Ethics, 2008). However, the nature of these issues varies depending on the perspective of the stakeholder involved. From the researcher’s perspective it is seen to be a waste of time, as limited resources are spent on document preparation for multiple ethics reviews. This is exacerbated by the fact that there is inconsistent feedback and associated requirements for revised materials from various sites, and even the potential for inconsistent REB decisions at various centres. This means that there is a potential impact on the integrity of research methodology and the validity of data resulting from substantially modifying studies, as required at local level as a recommendation of the REB. In addition, the time taken can lead to delays in the commencement of the research. Moreover, these delays have potential implications for the research participants and wider public on delaying research findings which could benefit the population (Interagency Advisory Panel and Secretariat on Research Ethics, 2008). From the public’s perspective,
the redundancy of reviews creates an unnecessary waste of public funds on human resources, wasted material resources and wasted time, and this potentially impacts on trust in the integrity of the ethics approval process in general (Interagency Advisory Panel and Secretariat on Research Ethics, 2008). Finally, from the REB’s perspective, the process constrains the ability of REBs where they have to rely on others’ decisions in multi-site trials; there is a lack of efficiency in the current process, and often duplication of reviews undertaken by local boards. In fact, a discussion paper on the ethics review of research in multiple settings and/or involving multiple REBs (Interagency Advisory Panel and Secretariat on Research Ethics, 2008) noted that constraining one REB’s ability to take action or make its own decision as a result of another REB’s decision was a key barrier to the successful approval of multi-site trials.

In response to these issues, an update was made to the TCPS, the tripartite council statement on research review processes (TCPS 2: 2nd edition of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans). As detailed in Section 3.1.3 above, this update allowed for additional options for review to be used in the case of multi-site trials. Briefly, these included: independent ethics review by several research ethics REBs; research ethics review delegated to an external specialised or multi-institutional REB; and reciprocal REB review.

Despite these options, it is unclear how widespread the uptake of reciprocal REBs or delegating to other REBs has been, and there is a limited literature describing researchers’ responses to the changes (Racine et al., 2011). For example, a recent 2012 publication opens with the sentence: ‘For ethical approval of a multicentre study in Canada, investigators must apply separately to individual REBs’ (Ezzat et al., 2010, p. 1). However, it does seem clear that REBs and their host institutions have been reluctant to enter into reciprocal agreements with other REBs, due to concerns about potential exposure to legal liability: they fear that review standards are not as strict in other locations (Glass, 2006). In addition, one source stated that since an institution is responsible for the research activities undertaken by its researchers, the institution that employs the researcher will review the research, irrespective of the location in which it is taking place. Therefore, it is not unexpected that duplication still exists within the system (Interagency Advisory Panel on Research Ethics 2009).

To alleviate the issues associated with relying on others’ decisions, ProGroup (a subgroup on procedural issues for the TCPS) suggests a series of measures to increase trust (Interagency Advisory Panel and Secretariat on Research Ethics, 2008). These include ongoing communication between central and local REBs, and transparency to aid the REBs to access previously completed reviews and engage at the local level. In order to facilitate this trust and transparency, the report suggests the public sharing of operational protocols by REBs. Another option detailed in the TCPS2 document is for official agreements to be drawn up between institutions, under which they will accept the decisions of each other’s REC’s review.

An evaluation of the limited guidance provided by the TCPS (Interagency Advisory Panel and Secretariat on Research Ethics, 2008) suggests several factors which may contribute to an effective review process when involving multiple institutions, and therefore review committees. These may include fostering trust between committees, ongoing connection
and communication between local and central committees, and buy-in of participating institutions, leading to a supportive policy environment and transparency within the system. Such transparency could be achieved through online public sharing of the institutional policies, procedures and application forms that they use. The above measures have been suggested through forums and discussions, and the successfulness of these will be visible through the system in the future.

3.5 Case study 3: The USA

As described in Section 3.1.7 above, the USA has a decentralised system which relies heavily on local review by IRBs at each institution where a study is being conducted. The literature suggests that one of the rationales for this decentralised review is to give appropriate attention to local concerns or issues and to protect local institutional liability (Fleischman, 2005). However, the extent to which this sentiment is borne out by empirical research into the function of IRBs is unclear, and some argue that this remains a significant gap in knowledge (Abbott and Grady, 2011). For one thing, alongside the benefits of local review, many barriers and challenges exist and there has been a sustained call in the research community for changes to the system, as observed through the present review of the literature. In a paper for the President’s Council on Bioethics (PCBE) in 2002, it was argued that there are up to 15 individual problems with the decentralised IRB system, which can be categorised under three different areas:

1. *problems in the structure of the human resource participants protection system* – including exclusive reliance on local, institution-based review; an absence of resources devoted to IRBs; inadequate education of investigators and IRB members; and repetition in the review process (particularly for multiple site studies)

2. *problems with the nature of the review process* – including a time-consuming review process, deficiencies in ongoing monitoring and excessive focus on informed consent forms


Others have summarised the issues in more general ways, including:

- inefficiencies of multi-centre trial review
- workloads and underfunding of institutional committees
- a lack of necessary ethical and scientific expertise among review members
- variability in the review
- an inability to manage individual and financial conflicts of interest (Mann and Shamoo, 2006; Abbott and Grady, 2011).

The complaints have not been heard only in the academic literature. A report by the deputy inspector general of DHHS in 1998 identified three key issues in relation to IRBs:
1. IRBs are unable to cope with rapid advances in biomedical research and changes in the research environment – therefore suggesting that structural changes to composition and membership are needed.

2. The failure, in some cases, to adequately involve the community in research decisions which might affect them – therefore suggesting that the ‘local’ nature of IRBs is not being used to best effect.

3. The problem of inadequate reviews due to increased workload – suggesting that proper resourcing of committees needed to be done (Grob, 1998).

A report for the Institute of Medicine a few years later also recommended that more integration between different types of review was needed, including in particular of scientific protocols (Institute of Medicine, 2002). This inadequacy of IRBs to address technological aspects of research was seen by many authors as particularly problematic (Mann and Shamoo, 2006).

The overly bureaucratic, time-consuming and costly nature of the review process is especially apparent when multi-site review is required. Estimates of the costs of IRBs vary, from 5,460 to 7,020 annual hours of senior investigator time at an institution hosting three IRBs (which is not uncommon at large research-intensive hospitals, universities or medical centres) (Association of American Medical Colleges, 2006), to an annual cost of US$76,726 for a small committee and US$289,416 for a large committee (Fitzgerald and Phillips, 2006). Others have found that the evidence for cost is too limited to be meaningful for meta-analysis; nevertheless, the evidence is sufficient to indicate there are cost burdens associated with the system (Silberman and Kahn, 2011).

Another source of concern in the literature is that of variation in the outcomes of assessment within different IRBs. This was repeatedly found in a systematic review of empirical studies of IRB practices (Abbott and Grady, 2011), and was particularly notable in approvals of multi-site research trials. In addition, there were problems found with interpretation of federal guidelines (Abbott and Grady, 2011). This raises questions about the nature of the inconsistencies and the implications for research. Some claim that the differences in IRB review can jeopardise the scientific integrity of the research, as different changes are requested to meet local needs. Many suggest that more guidance from federal institutions could be one way of addressing this problem.

In light of these issues, calls for a more centralised, or at least reformed, system have become a recurrent theme. The benefits of a centralised system would be expert review specific to the research at hand, consistency in the protection of human participants, a reduced burden on local IRBs, and potentially reduced time spent obtaining approval (McWilliams, Hebden and Gilpin, 2006). One author points out that in order for a national and local system to work in concert, there needs to be reform to central legislation, better guidance at a national level, improved resourcing of local IRBs and a greater awareness of researchers as the ‘customer’ of IRB functions (Fleischman, 2005).

In order to achieve a more streamlined system, many have proposed that different elements could be considered, such as:

- Introducing a certification process for people on IRB panels (indeed, this is a feature of the centralised review processes discussed below and the role of the
Association for the Accreditation of Human Research Protection Programs – a national accreditation body

- providing leadership and guidance at a national level
- introducing information management systems to streamline the exchange of information
- providing greater resources to IRBs, and
- encouraging national advisory bodies to issue guidelines on good practice.  

In response, changes to the system of IRB review are under consideration by the DHHS which would both improve the effectiveness of IRBs and enhance human participant protection (Ezekiel and Menikoff, 2011; US Department of Health & Human Services, undated). The most significant of these within the context of this study are proposed changes to the system of review for multi-site studies, and improvements to the mechanisms through which guidance is provided by different federal agencies with jurisdiction over different types of research studies and trials.

Another important development has been the establishment of two major CIRBs within the US Department of Veterans’ Affairs and the National Cancer Institute. These committees allow for a specialised composition of committees, with topic and methodological expertise. This more centralised type of system has advantages and may help to allay some of the concerns addressed in the literature about the lack of appropriate expertise on all ethical review boards.

The model of a central IRB was developed initially by the National Cancer Initiative (NCI) in collaboration with the Office for Human Research Protections at the DHHS, as part of an initiative to improve the NCI’s clinical trials programme (Adler, 2011). It was introduced first in the mid-2000s. The NCI adopts a shared responsibility model, where a local IRB is used to understand local context and factors which might affect the research, but this local IRB signs off all oversight and review authority to the CIRB. In this set-up, the CIRB’s primary function is initial and continuing review of the trial, while the local IRB is responsible for conduct of the trial, local standard operating procedures and locally adverse occurring events. The local IRB also participates in the initial review through a ‘facilitated’ process, whereby it is able to input to the central review by advising on local context considerations.

Another solution implemented in 2011 is the use of IRBShare, an online service which shares review documents and procedures. This software aims to build trust and confidence between review boards. Currently the initiative is being piloted in support of New York University’s multi-centre trial.

However, issues of trust are observed between IRBs when dealing with centralisation initiatives. The NCI has reported that although many institutions may sign up to

---

11 See for example Fleischman (2005) and McWilliams, Hebden and Gilpin (2006).
facilitated review, they still require local investigators to go through the local IRB process, creating extreme duplication. An interviewee is quoted in one peer-reviewed publication thus:

I like the decentralization, not because I like repeating all the work, but because I’m not sure I would trust a central IRB somewhere in Washington. (Klitzman, 2011, p. 7)

The NCI is working with local institutions to convince them of the benefits of centralised review, but it seems that this is a process that will take time for trust to develop and will need to rely on the continued supply of data to show the benefits.

3.6 Conclusion

The challenges of research governance are not necessarily unique to any country, but the mechanisms through which people try to address the problems do vary in their success and implementation. Nevertheless, the present analysis has identified common barriers and ways of overcoming them which are summarised here.

Due to the inherent risk to human health associated with the decisions of ethics committees, it is important to clearly and publically establish their role and authority. For example, developing clear guidance on the roles and responsibilities is crucial: in particular, developing an understanding of what the committee and institution’s legal liability is if they accept other’s review decision, or if there is an issue with the study. This could be communicated through training and educating the various groups involved in conducting and regulating research, to enable better management of the actual risk.

Australia uses a system of institutional accreditation to instil trust in review boards which receive decisions from other review boards. The requirement of organisations to meet accreditation measures, set by the director-general of New South Wales Ministry of Health, gives a level of responsibility and authority to the subset of review boards conducting reviews on behalf of others. Alternatively, many have commented that the US system could use certification of staff (Levine, 2011) to provide the required mutual trust in others’ decision making. The establishment of national standards or guidance for this training, such as that provided by the Association for the Accreditation of Human Research Protection Programs, could help to provide confidence that everyone has received the same level of training to make these decisions.

Agreements are another way of establishing trust and confidence in other committees’ decisions. These could be set up between individual committees under which they will accept, with an agreed level of oversight, the research ethics reviews of each other’s review committee. This is used in the Canadian Reciprocal REB Review, and also seen through the memoranda of understanding signed between states in Australia. This officially implies and declares a level of trust between establishments that they have an agreement to accept the reviews of others.

In the UK, site-specific assessment occurs simultaneously with overall review by the centralised NRES. NHS Trusts do not have the approval information from NRES before they conduct their own approval, and therefore some replicate the work that NRES is undertaking. In Australia, these processes occur sequentially, allowing each person to be assured of the approval before them. This sequential approval also happens in Brazil,
although it has been pointed out as a source of delay within their system. China seems to overcome this by establishing strict timelines for sign-off on projects. Consideration of whether sequential, step-wise processes could reduce duplication of review is worth considering, although it does seem to be something that varies by country as to its success.

A mechanism suggested to improve the use of multi-site REBs within the Canadian system is to foster a culture of trust through greater transparency of the process. An overall culture of openness encourages trust, as people feel that they are part of the bigger process and that the whole process is accessible to them. The culture of openness that transparency creates is fundamental to building trust between parties. Therefore, within the research review process, if committees were able to see the full spectrum of information that others had used to reach a decision, this would help in trusting others’ conclusions.

In addition to this, face-to-face interactions and establishment of trust over time have been suggested as important when centralisation or reciprocal review processes are introduced. Increasing relationships between committees could remove the facelessness of another institution. These interactions, in turn, could create trust, as individuals could picture the person who is taking the decision on their behalf, rather than relying on someone they do not know. Therefore, increasing interaction and communication between committees could be a mechanism to increase trust and confidence in other committees (Hedgecoe, 2012).

Stakeholders quote anecdotal examples of where things have gone wrong or near misses as a reason to duplicate reviews. The frequency of these events is minimal, but these extreme examples can be fixated on; therefore, it is important that those involved in review understand the relative risk. As stated previously, in Australia there have been very few cases where the researcher may be sued, but RECs are rarely, if ever, sued. This trend is also seen in the USA and the UK, where, we understand, the NHS Litigation Authority has received no claims relating to research. A conference in the UK concluded that in the absence of any direct contact between a REC and the researcher, there is no basis for care beyond what is normal and reasonable in carrying out their limited advisory role. Commentators in the USA suggest that the number of problems with research trials only receive so much press and media attention because they are so rare (Association of American Medical Colleges, 2006). In addition, a systematic review of US IRB empirical studies does suggest that we have no clear understanding of how to evaluate IRB outcomes in standardised ways which can enable joint learning and improvement. Therefore, an evaluation or audit of the system to determine the overall success of research could produce confidence in the current system. This information could be fed into the accreditation scheme described above, to ensure that a certain standard and success rate is seen at institutions providing recommendations to other institutions.

Finally, Australia is using standard operating procedures: this method is intended to reduce duplication by reducing uncertainty in others’ decisions. It applies a standard set of criteria to all reviews, creating confidence that others have applied the same rigorous process to approval. As the legal liability for ethical review remains with the institution, there can be a lack of trust that other reviewers are conducting their review to the same standard.

---

13 This emerged from a conversation with the DH in the scoping phase for this project.
Therefore, often individuals would rather replicate work assessing approvals already granted, than sign off on the strength of someone else’s opinion.
Chapter summary

This chapter summarises the regulatory responses to certain key events in the medical drug industry in the UK and the USA over the postwar period. Specifically, it considers the Thalidomide and AIDS crises, the introduction of European directives, and the establishment of the European Medicines Agency. These events provide good natural experiments for understanding the effect of policy changes on the regulated sector, as the immediate changes in the regulatory stance provide a benchmark against which changes in behaviour can be examined.

Based on the present review, it is suggested that the ways in which changes to the regulatory system for medical drugs are received and responded to by stakeholders offer lessons for the regulation of health research. These include the following.

- Manage seemingly inconsistent regulatory objectives carefully so that they are complementary and not contradictory. Failing to properly balance objectives may result in a cycle of increasing and relaxing standards, eroding confidence in the regulator. Further, if the result is not fed through the regulatory body to the relevant executing units, they may experience delays in conducting reviews.

- Develop clear, consistent guidelines on what constitutes a poor outcome (such as unnecessary risk to human lives) and how these will be handled, including setting out, where relevant, criminal and civil liabilities for non-compliance. Guidelines set out an ex-ante set of expected punishments, reducing the likelihood of ad hoc, post-crisis pronouncements and ensuring that public trust in the system is maintained.

- Leverage public pressure to enhance industry compliance with regulations. By educating the public on what it considers to be acceptable standards, the regulator can engage the public as an alternative compliance monitor.

- Consider regulatory actions in the context of the overall landscape in which firms operate: for example, return on investment in drug development is affected heavily by developments in the intellectual property rights landscape. Regulators must consider the likely impact of new regulatory actions on firms’ ability to extract profit under intellectual property rights, in order to understand and/or predict compliance behaviours better.
4.1 Introduction

The regulation of medical drugs is an important element of the health system. The actions of regulators in this sector have a direct and often immediate impact on patient health, and as such tend to be highly event-driven. Indeed, all important legislative changes observed in this review within the USA have been in reaction to specific events. However, in the UK the regulatory landscape has been shaped more by European directives than by specific events, especially in recent years. This chapter explores the regulatory responses in medical drug approval undertaken after key events, as the immediate changes in the regulatory stance provide a benchmark against which changes in behaviour can be examined. It then offers a brief overview of regulatory approaches toward pharmacovigilance, before presenting some key lessons from medical drug regulation.

4.2 Key regulatory events in US medical drug approval

The US Food and Drug Administration (FDA), which sits within the US Department of Health & Human Services, is responsible for ensuring the safety and efficacy of drugs used on humans or animals, for approving the manufacture, marketing and distribution of all such drugs, and for monitoring the effects of these during the period of sale. More broadly, the agency’s responsibilities extend to advancing public health, and on these grounds the FDA regulates the marketing and distribution of biological products, medical devices, food, cosmetics, tobacco products and any products that emit radiation.

Typically, the current approval process for pharmaceuticals is described by phases. Generally, once the applicant company has completed the laboratory and, where relevant, animal tests and is ready to begin testing in humans, it submits an Investigational New Drug application to the FDA. This allows the company to begin clinical trials with a small number of healthy volunteers (Phase I trials). Conditional on an appropriately low level of adverse reactions being observed, the company may proceed to Phase II, when the drug is tested on a small number of patients suffering from the relevant condition, in order to determine the efficacy of the product and again verify the safety of the product. If this stage is successful, the company proceeds to Phase III, where the drug is trialled across a large number of both healthy and affected volunteers. The applicant must conduct at least one blind study (a study where the patients do not know if the drug they are given is the product being tested or a placebo), although generally it is thought that inclusion of a double-blind study (where neither the patients nor the medical professionals administering the doses know if the patient is receiving the product being tested or a placebo) is a sign of high-quality testing. Once the studies are concluded and it is determined that the drug is likely to meet the criteria for marketing approval, the applicant submits a New Drug Application, which includes the major results of the trials as well as manufacturing and packaging specifications, along with any laboratory or animal studies not included at the Investigational New Drug stage. The FDA then reviews this application and decides whether to grant marketing approval.

Where drugs are derived from biological material, the last stage of the approval is different. Instead of a New Drug Application, the applicant seeks a biologic licence application from either the FDA’s Center for Drug Evaluation and Research, or the Center for Biologics
Evaluation and Research. The entire process can take in excess of one year to be completed: for the period 1993–2003, the FDA reports that median time to approval for drugs via the standard track\(^{14}\) was 16.2 months.

### 4.2.1 Specific event: Thalidomide

**Learning from a lucky escape**

Thalidomide was developed in Germany in the immediate postwar period: it showed promise as an anti-anxiety drug and as an aid to moderate the effects of morning sickness during pregnancy. It went on sale in the then-West Germany in 1956 and was widely available throughout the continent within a few years. However, shortly after widespread prescription began, reports of serious birth defects that could be associated with the drug started to become public. Given that the drug sponsors did not submit an application for marketing in the USA until 1960, American policymakers were able to learn from these reports. Indeed, the FDA engaged in strategic delay of the application until the reports could be substantiated, and in fact never approved the drug until decades later (see below).

However the USA did not completely avoid the adverse effects of this product. At that time, the version of the Federal Food, Drug and Cosmetic Act of 1938 under which the FDA was operating did not require that clinical trials receive FDA approval. As such, Thalidomide trials, intended to prove safety of the drug, were able to begin in the USA in 1958. It is estimated that more than 20,000 people, including some 624 pregnant women, were exposed to the drug (McGrath, 2005).

While the USA did not suffer the widespread effects seen elsewhere, the effects on the limited number of trial participants, as well as the huge public outcry overseas, did spur policymakers on to reform the existing legislation. The Act was not only deficient in omitting the supervision of clinical approval, but it also required that the applicant only show the safety, not the efficacy, of a drug for approval purposes. It gave the FDA only 60 days within which to act on an application – if no objection was made by this deadline, marketing could proceed automatically. These aspects of the law were all amended in the Kefauver-Harris Amendments (also called the Drug Amendments) of 1962, which set the framework of the current system to require FDA supervision of clinical trials, incorporated the efficacy requirement, and extended the time for the FDA to review an application to 180 days. The Amendments also made ‘affirmative approval’ by the FDA a requirement, so that companies could no longer presume automatic approval if no response to their application was forthcoming.

The Amendments served to calm public (and some bureaucratic) fears that regulation was too lax, leaving the USA vulnerable to events such as the Thalidomide crisis. However, almost as soon as they were passed, industry countered with the suggestion that the regulatory response had been too extreme, and would lead to a slowdown in delivery of new drugs to the market. Indeed, there is some evidence that in raising the hurdles for

---

\(^{14}\) The standard track is the usual approval process for medical drugs. Sponsors may apply for accelerated approval (discussed further in section 4.2.2) if they wish to bypass the standard track.
approval, the legislation effectively led to an increase in the delay of new drug approvals\textsuperscript{15} and an increase in the cost of pharmaceutical R&D.\textsuperscript{16} As these outcomes became evident, drug companies and congressional dissenters began lobbying for further legislative action to moderate these effects. Among the actions subsequently taken was the Drug Price Competition and Patent Term Restoration Act of 1984, which allowed pharmaceutical companies to claim up to five years of additional patent protection to cover the extended time that drugs spent awaiting approval. The Act also removed the FDA requirement for independent testing of generic versions of brand name products, allowing those generics to reach the market more quickly. This regulatory ‘cycling’ between increasing standards on safety grounds, and relaxing them on development grounds, may be a result of the event-driven nature of the industry, as reports of adverse effects among approved products prompt internal and public demand for strict regulatory action, while reports of drug shortages or undertreated conditions drive demand for relaxed regulation.

**A second look at high-risk drugs**

While Thalidomide clearly had been ruled out for widespread prescription in the population, it remained extremely effective in treating certain serious conditions that were non-responsive to other drugs. Among these serious conditions are leprosy (Thalidomide remains the only product effective against the erythema nodosum leprosum component of the disease) and HIV/AIDS (Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome: Thalidomide is effective in reducing cutaneous and oral lesions as well as the wasting associated with the condition) (McGrath, 2005).\textsuperscript{17}

Therefore, the FDA faced a dilemma: it could not ban the drug completely, as it would seriously affect those suffering from the aforementioned conditions and could lead to black market activity (the drug remained legal in many other countries, so smuggling could develop), but neither could it simply approve the drug only for the treatment of specific diseases, as existing regulations allowed physicians to prescribe drugs for off-label uses. As neither the approval nor prohibition of Thalidomide could be expected to produce the desired outcome, the FDA had to develop a new mechanism to induce responsible, targeted use of the product by the sector. The result was an innovative new monitoring system, which allowed the agency to approve high-risk drugs while maintaining relatively tight control over drug distribution. The System for Thalidomide Education and Prescribing Safety Program mandates the registration of the doctors prescribing as well as the patients receiving the drug, and imposes strict behavioural and reporting requirements

\textsuperscript{15} One estimate is that the Amendments increased the drug development process from one to four years to seven to 13 years (see Roberts and Bodenheimer, 1982).

\textsuperscript{16} Estimates of the effective cost of the Amendments to the US economy are contested. Scholars agree that there were substantial increases in development costs over the decade from the mid-1950s to mid-1960s, and that the Amendments were a significant factor, but they differ in the approach to disentangling the effects of the Amendments from general input cost increases (see Schifrin, 1982).

\textsuperscript{17} Further, as late as 2003, the drug was also being trialled for treatment of neurofibromatosis (‘Elephant Man’s’ disease). It is now estimated that thalidomide may be useful in the treatment of up to 130 conditions.
on both sets of parties (McGrath, 2005). This strict monitoring system allowed the FDA to finally approve Thalidomide in 1998, and has provided a novel approach to cost–benefit considerations for high-risk drugs, offering a guide for FDA treatment of other such products.

4.2.2 **Specific event: the AIDS crisis**

The 1980s saw the outbreak of AIDS, a fatal infectious disease for which conventional treatments were ineffective. As the epidemic took hold, the delay in drug approval in the USA relative to Western Europe (highlighted in the wake of the 1962 Drug Amendments) became the focus of serious public debate. Further, the FDA’s traditional argument that delays should be considered the ‘price to pay’ for safety was undermined, as patients were dying without treatment.

There was increasing pressure from AIDS activists for patients to access drugs which had cleared Phase II trials already. This was particularly relevant in respect of azidothymidine, which was the first compound to be found effective in Phase II trials against HIV, which causes AIDS. Under pressure from the increasingly vocal AIDS lobby, as well as an increasing number of members of the public, the FDA agreed to accelerate the review process for azidothymidine and approved the drug within two years of the Investigational New Drug application (Greenberg, 1999).

In addition to this, in 1992 the FDA introduced new rules, the Accelerated Approval (Subpart H) Regulations, which accelerate approval for drugs to treat life-threatening and serious diseases. These rules also enlarged access to pre-approval drugs for patients with limited alternative treatment options. This so-called ‘parallel track’ meant that while Investigational New Drug approvals remained a requirement for all new drugs, those deemed appropriate for treating HIV/AIDS were put through an abbreviated testing phase, where appropriate ‘surrogate’ endpoints were applied.

In the same year, Congress passed the Prescription Drug User Fee Act of 1992, which required drug and biologic sponsors to pay a fee when submitting the product application. A portion of this fee was then used by the FDA to hire more reviewers, with a view to expediting the approval process.

Concerns about the robustness of the ‘rushed’ trials and calls by other constituencies for the FDA to move away from ad hoc decisions – which often involved prioritising applications within the FDA, refocusing scarce agency resources away from other products and toward a systemic reform to accommodate accelerated treatments for terminal conditions – led to a series of policy refinements. The Food and Drug Administration Modernization and Accountability Act of 1997 not only formalised the accelerated approval process, but also made ‘prompt and efficient review’ an explicit mission of the FDA (Greenberg, 1999, p. 345). This need for clarity and consistency across serious conditions suggests that while the policy response was event-driven, perhaps it involved a too-specific target.

---

18 McGrath reports that ‘doctors… must register… and ensure that their patients fully understand the risks involved’ and that ‘patients must practice two forms of birth control… and female patients of child-bearing age must undergo monthly pregnancy tests’ (2005, p. 618).
4.3 **Key regulatory events in UK medical drug approval**

4.3.1 **Specific event: Thalidomide crisis**

Until the Thalidomide crisis of the 1960s, drug approval in the UK was essentially unregulated. There were no industry-wide safety or efficacy standards, and the industry was constrained only by fraud laws, which required truth in labelling (Teal, 1985). However, after the crisis, parliament passed the Medicines Act 1968, which not only introduced safety and efficacy rules for approval, but also instituted specific classifications (prescription, pharmacy-only and general) for medical drugs.

In the wake of this legislation, drug review was conducted by the Medicines Division of the DH. This division became the Medicines Control Agency in 1989 and merged with its medical devices counterpart, the Medical Devices Directorate, to become the Medicines and Healthcare Products Regulatory Agency (MHRA) in 2003. As currently constituted, the MHRA, an agency of the UK Department of Health, is the body authorised to approve medical drugs and devices for sale in the UK and to monitor their effects while on the market.

The MHRA typically has only three types of interaction with researchers and/or sponsors during the approval process for a medical drug. The first, scientific advice, is available any time before sponsors seek marketing authorisation, but usually is sought in the early stages of clinical trial preparation, when sponsors are finalising trial protocols. However, as the MHRA does not provide any overall review of the product at this stage, the sponsor must approach the Agency with specific questions about the trial design. This step provides sponsors with information about what steps – with specific relevance to the questions posed – would be appropriate for researchers to take in order to conduct a trial that will have a high probability of leading to marketing authorisation for the product. Scientific advice is an entirely optional step, and the feedback that the MHRA gives is not legally binding on the researchers, although there is evidence at the European level that products for which sponsors seek and follow the recommendations of scientific advice have a higher probability of receiving marketing approval (Regnstrom et al., 2010).

The second is in approval for clinical trials. The MHRA authorises the conducting of clinical trials at UK sites, although legislation restricts its focus to issues of patient safety. The third is the most well-known step, marketing authorisation approval. Regulations limit the MHRA to three possible grounds for refusing authorisation:

1. where the quality of the product is not as advised
2. where the benefit-to-risk ratio associated with product use is adverse
3. where the product offers no therapeutic efficacy.

4.3.2 **Specific event: European directives and the European Medicines Agency**

The last 30 years have not only been a period of significant organisational change for the national regulator, but have seen significant changes in regulation itself, as UK rules have been largely harmonised to and/or superseded by European directives on medicine and medical devices.

The European Medicines Agency (EMEA) was established pursuant to Council Regulation (EEC) no. 2309/93 in 1993, and works with national agencies to coordinate the roll-out of
directives throughout the EU. Most relevant here are Directives 2001/82/EC\textsuperscript{19} and 2001/83/EC,\textsuperscript{20} which set out the guidelines for the regulation of human and veterinary drugs and devices in the EU. The Agency works via seven committees, although the committee relevant for medical drug approval is the Committee for Medicinal Products for Human Use.\textsuperscript{21} Via this committee, the EMEA provides a centralised EU approval process for drugs and medicines as an alternative to national approval. This is now the compulsory approval process for drugs for selected important diseases including HIV/AIDS, cancer and diabetes, as well as for drugs for orphan (rare) diseases and medicines from biologics. The EMEA also facilitates the ’mutual recognition’ process, whereby approvals of one Member State are recognised by the other Member States.

However, the EMEA functions effectively as a secretariat. Members of the Committee for Medicinal Products for Human Use are drawn by appointment from national authorisation agencies. Any evaluation of a proposed product via the Committee is led by a committee member who is elected rapporte:\textsuperscript{22} in practice, usually this can be interpreted to mean not the individual member, but the national agency at which the rapporte\textsuperscript{r} is based, as all the analysis for the authorisation is conducted there. This national agency also retains primary responsibility for the product post-authorisation.

As a result, coordination is an important, ongoing issue both within the EMEA and between the EMEA and national agencies such as the MHRA. Not only is EMEA (like other EU agencies) required to operate in a way that is effective for countries with disparate geographical, cultural and economic characteristics, but in this specific area, it also must standardise approvals carried out under different rapporteurs at (or in line with) different national agencies. Furthermore, approval decisions are complicated by the fact that drug classifications and drug distribution systems\textsuperscript{22} differ across Member States.

The indications are that this centralisation is proving successful in stimulating faster diffusion of medical drugs throughout the EU (Varol, Costa-Font and McGuire, 2012), although it has introduced distortions in respect of the timing of product launches. National pricing policies differ across the EU, opening up the possibility of ‘parallel trade’, where drugs may be cross-sold from a low-price Member State to a high-price Member State. As a result, while companies are making increasing use of the centralised approval mechanism, they seem to be strategically timing the sale of products, launching first in high-price locations and delaying marketing in low-price jurisdictions (Varol, Costa-Font and McGuire, 2012) to minimise arbitrage opportunities. This may have implications for public health and well-being in the latter Member States and for equity across the EU.

\textsuperscript{19} As amended by Directive 2004/28/EC.

\textsuperscript{20} As amended by Directives 2004/24/EC and 2004/27/EC.

\textsuperscript{21} The other six constituent Committees of the EMEA are the Pharmacovigilance Risk Assessment Committee, the Committee for Medicinal Products for Veterinary Use, the Committee for Orphan Medicinal Products, the Committee on Herbal Medicinal Products, the Paediatric Committee and the Committee for Advanced Therapies.

\textsuperscript{22} Certain Member States have national price controls and/or formularies.
4.4 Regulatory attitudes to pharmacovigilance

Approval is but one component of medical drug regulation. Post-approval, national agencies also must monitor the effects of the products on individuals during the period of sale. This function, termed pharmacovigilance, is critical, as while approval is on the basis of sound clinical testing, national (or EU-wide) marketing exposes drugs to a larger population with much more diverse health characteristics and a broader set of prescribed products with which the drug may react. Thus, pre-market tests alone are insufficient to categorise a drug as safe or unsafe.

Much like the regulation of drug approval, pharmacovigilance regulation tends to be driven by individual events. As mentioned previously, in the USA the drug regulation system had been structured by the Federal Food, Drug and Cosmetic Act of 1938, and was made more formal by the Drug Amendments of 1962. The Thalidomide crisis as well as a series of drug withdrawals in the 1970s and 1980s meant that more often than not, the FDA aimed for the side of caution. By 1999, the FDA had a well-developed ‘risk management framework’ for drug safety (Demortain, 2008), which has underpinned a series of drug recalls over the past decade. Indeed, between 1997 and 2005 the FDA withdrew 18 drugs from the market, compared with only nine over the preceding 20 years (David, 2009).

However, in the UK the system was much more informal, with regulation only having been introduced in the late 1960s, although this relatively more relaxed approach became more formal over time with the introduction of European directives. A key turning point for EU guidelines on pharmacovigilance came with the withdrawal of Cerivastatin, a blockbuster drug approved to lower lipids which eventually led to more than 100 deaths before being withdrawn from the market in 2001 (Demortain, 2008). This drove the EMEA to set EU-wide guidelines on pharmacovigilance, and spurred the international adoption of pharmacovigilance planning guidelines (Davis and Abraham, 2011).

Over time, repeated refinements and strengthening of EU regulations have meant that the European regulatory environment is now more formal than the US environment – a reversal of the state of play in the 1970s and 1980s which is often termed the ‘flip-flop’ hypothesis (Davis and Abraham, 2011). In spite of some degree of collaboration between the FDA and the EMEA in recent years, the difference between the systems remains, with the FDA seeming to be more willing to engage in risk management of potentially problematic drugs on the market, while the EMEA seems to respond to potential issues by withdrawing drugs from the market (Davis and Abraham, 2011).

---

23 Among the more noteworthy of these recalls were the diet drug Phenylpropanolamine (recalled in 2000), and Merck’s former flagship arthritis drug, Vioxx (recalled in 2004).

24 In 2003, the EMEA and FDA concluded a confidentiality agreement which allowed them to share information on marketing authorisation processes as well as post-approval surveillance. The agencies also offer joint scientific advice (where sponsors so request), allow single submission of annual reports for orphan (rare) disease products, and collaborate on inspections relating to the international Good Clinical Practice Initiative.
4.5 Lessons from medical drug regulation

There are several key lessons to be drawn from the events presented above. First, while regulation in all sectors is driven to a certain extent by episodes of crisis, medical drug regulation appears to be especially sensitive to individual events. Indeed, approval of a single product which proves dangerous is enough to shake public confidence in the regulator. This drives national agencies toward strict approval and monitoring criteria. However, strict criteria mean higher development costs, longer testing regimes or greater reporting requirements for pharmaceutical firms – all of which raise fixed costs and may negatively affect investment and/or product innovation. Thus, the regulator must manage the seemingly juxtaposed, yet potentially complementary, objectives of ensuring drug safety while simultaneously stimulating investment in new products. There is a parallel here to consider in the Health Research Authority’s own aims of ensuring high-quality, ethical research while also stimulating clinical trial activity in the UK. The two can be complementary, but this is the challenge of regulation and research governance. It appears that in the 1980s and 1990s, US agencies were under sustained pressure to focus on availability as a result of the AIDS crisis (an excellent illustration of the power of well-organised lobbies) until there were important drug withdrawal events, when their priority switched back to safety. It is important that priorities are set nationally and clear guidelines set out for how to implement them. This would be particularly important in the case of research governance, where the priorities decided at an appropriately central level should be fed down to RECs in a clear fashion, so that the units themselves do not try to balance competing objectives.

Second, policymakers must develop clear, consistent guidelines on how specific poor outcomes will be handled: including setting out, where relevant, criminal and civil liabilities for non-compliance. In this way, regulators can maintain public trust in the system while minimising the reputational effects associated with the ‘blame game’ often played out before the cameras in the aftermath of a crisis.

Third, firms in the industry face contradictory but potentially complementary objectives: on the one hand, they are attempting to maximise the private return on investment, while on the other, maintaining a healthy public image. As such, by educating the public on what it considers to be acceptable standards in drug development and delivery, regulators can leverage public pressure to enhance industry compliance with regulations.

Fourth, policymakers should consider action in the context of the overall landscape in which firms operate: specifically, return on investment in the pharmaceutical industry is affected heavily by developments in the intellectual property rights landscape. We have seen in the USA that raising approval standards was matched by an extension of patent terms to keep drug R&D relatively steady, via the Drug Price Competition and Patent Term Restoration Act of 1984. Careful consideration of changing trends in the intellectual property rights or other relevant landscape will help policymakers to understand fully and/or predict more accurately the likely responses of the sector to specific regulatory changes.
Chapter summary
This chapter summarises the insights to be gained from environmental risk regulation. First, it discusses overarching issues in the sector and trends in regulation, including a broader move towards a more dispersed model of environmental governance. Second, it discusses specific case study examples in different areas of environmental risk and regulation, including contaminated land management, risk-based decision making and trends in corporate environmental management. One of the core issues underpinning environmental governance and driving changes in the sector is the tension between the economic growth of companies and environmental and health risks to the public. It is the role of the regulator to protect the latter, but it is also in the wider interest of governments to encourage and stimulate economic growth. Mechanisms and issues affecting the way in which the regulatory system is received and responded to by stakeholders include the following.

- Harness the role of public trust and confidence – demand from the public for environmental accountability was not only an early driver of regulatory action, but is also a current driver of proactive corporate environmental management.
- Equally, harness consumer demand, particularly where the government has less direct control over the behaviour of the regulated sector.
- Ensure that there is no misalignment of regulatory and actor philosophies, which can pose a threat to implementation of regulation, and hence present challenges for effective and efficient governance responses.
- Use education, training and capacity-building to encourage actors to engage with each other and to foster understanding of the views of different stakeholders, in order to create a system that is seen by all as more legitimate and effective.
- Use incentives as the actors become more dispersed and behaviours more difficult to control. Here it is necessary to take into account the motivations of different actors, and shape incentives accordingly.
Since the 1960s, there has been a growing awareness of the tension between industrial and technological advances and environmental risk. As industry develops more advanced technologies to meet the needs of modern economies, the impacts of these technologies on the environment, and often on public health, have risen in the public and regulatory consciousness. As a result, environmental risk management has moved up the regulatory agenda.

This chapter summarises the insights to be gained in relation to the shifting nature of environmental risk and regulation, the corporate governance responses which have emerged and the role of external stakeholders, in particular the public, in the wider system. First, it discusses the overarching issues in the sector and trends in regulation, including a broader move towards a more dispersed model of environmental governance. Then it discusses specific case study examples in different areas of environmental risk and regulation, including contaminated land management, risk-based decision making and trends in corporate environmental management.

5.1 The rise of environmental governance

Environmental legislation covers many different areas from pollution to radioactive waste, but shares in common the idea that we should protect the environment and human health through regulatory oversight, while still allowing for economic growth (for example, not stifling industry). The so-called ‘toolkit’ of regulatory instruments most commonly seen in environmental regulation includes: ‘command and control’; economic instruments; information-based instruments; co-regulation and self-regulation; support mechanisms; and capacity building (Irvin, 2002; Taylor et al., 2012). Command-and-control mechanisms are most useful when there is a clear source of an environmental pollutant and it can be directly targeted by a regulatory measure. Some argue that these kinds of direct regulations stifle innovation, as they force corporations to conform to exact standards or practice. Historically, in the USA the Environmental Protection Agency has tended to take this much more inflexible, command-and-control approach to environmental regulation (Rondinelli and Berry, 2000). This approach has been criticised as leading to reduced environmental performance. However, there is some interesting evidence about command and control as an example of stringent regulatory controls forcing operators to innovate. The so-called ‘Porter Hypothesis’ is an active area of debate, but one that is intriguing for its stance on the positive effects of regulatory activity (Taylor et al., 2012).

Conversely, economic instruments operate by changing the incentives faced by firms or individuals to encourage them to voluntarily change their behaviour. Other options in the same realm are payment schemes which pay to recognise good behaviour, and there is

25 It is worth noting the parallels here between these regulatory instruments and those which are used in other sectors. In particular, Braithwaite, Healy and Dwan (2005) identify similar tools in relation to healthcare regulation: they classify them according to whether they are ‘hard’ or ‘soft’ instruments and analyse their utility within broader frameworks of ‘responsive regulation’, ‘networked governance’, ‘meta-regulation’ and ‘restorative justice’.
some evidence to show that this is effective in encouraging positive behaviours in ecosystem management (stewardship management) (Taylor et al., 2012).

Information-based instruments are a way to provide better evidence on which to make decisions. Naming and shaming, targeted information provision and registration, labelling or certification schemes are all different types of instruments which could be used. There is little evidence for improved environmental performance from naming and shaming, but there is some evidence of a negative impact on businesses’ performance because of reputational losses. Many labelling schemes have been established now for environmental standards: measures on how this actually impacts the environment are difficult to assess, although they probably are having an effect on ecosystem management. One study provides evidence that ecological certification schemes can help consumers play a role in environmental regulation, which can be an important driver of behaviour and improved environmental performance within the sector (Irvin, 2002).

Co-regulation or self-regulation can allow businesses to identify the best ways to operate, but also can carry great risk. However, the evidence is mixed, and some suggest that this approach may be effective only when accompanied by the threat of stricter regulation. This speaks to the importance of developing shared understandings of the environmental outcomes against which to measure. Finally, support and capacity building can be beneficial, but research evaluations have shown that this is only effective if the programmes are targeted appropriately to the context and industrial sector (Taylor et al., 2012).

These instruments have all been used over the past 20 years as environmental risk awareness has risen in the eyes of the public. Major environmental catastrophes, such as the Love Canal toxic waste disaster in the USA and the Chernobyl nuclear disaster in the former USSR, have made the public and governments acutely aware of the human health risks posed by environmental damage (see for example, Jasanoff, 1995 for a history of many of these). As a result, the public has become more demanding of regulators to protect the environment for future generations. Conversely, there are pressures from corporations and industry: this group is more likely to resist the rising costs associated with increased regulation and regulatory compliance, as these costs are damaging to their economic bottom line. These two drivers of behaviour within the regulatory system are in tension, but not at odds. Governments need to encourage and enable risk-taking behaviour for growth while simultaneously curbing or limiting environmental damage (Gouldson and Bebbington, 2007). As a result, it is argued that state capacities to intervene have been reduced, leading to the introduction of different types of mechanisms and models for regulation and environmental governance.

In particular, there has been a movement from direct regulation based on mandatory operating requirements, inspections or enforcements – which are good for targeted problems – to a regulatory model that

seeks to harness other social forces such as the buying behaviour of consumers or customer–supplier relationships amongst businesses, to influence business and individual behaviour. (Taylor, 2012, p. 271)

All of this has led to the emergence of a concept of environmental governance or ‘decentred’ risk regulation, where authority is dispersed and governance becomes a ‘multilevel, multiactor phenomenon which is complex and fragmented, with new patterns
of interaction and great variety’ (Gouldson and Bebbington, 2007, p. 7). In this, there is a greater emphasis on self-regulation through corporate social responsibility initiatives; a greater opportunity for market and civic actors to play a role in governance; and opportunities for governments to create new modes of participation that move beyond solely technocratic, scientific approaches to decision making. The next section looks at different examples where this more dispersed approach is apparent.

5.2 Environmental governance and contaminated land management in the UK

The issue of land contamination did not seep into the national and international consciousness until the 1960s and 1970s, with several high-profile incidents attracting major media attention and prompting politicians to take immediate action to try and remove all pollution or threat of pollution entirely (Ferguson, 1999). However, elimination of all risks is neither technically nor economically feasible, and so generally speaking, globally the focus has shifted from cleaning up all sites to managing contaminated land and making it fit for future use (Ferguson, 1999).

In the UK, it has been estimated that there may be anywhere from 50,000 to 200,000 hectares of contaminated land (Royal Commission on Environmental Pollution, 1996), with an associated cost of £20 billion to £40 billion to regenerate and remediate it (Watson, 1993). Currently, in the UK contaminated land is dealt with through a risk-based approach to regulation and governance. Like much of environmental regulation, the approach seeks a balance between ensuring the protection of public health and ecological systems, alongside promoting the regeneration of polluted sites, while keeping costs down for the companies involved (the regulated sector). Large-scale remediation in the UK began in the 1960s with the regeneration of sites such as the Lower Swansea Valley, but there was no formal government institution in place to oversee these processes until 1976, when the Inter-departmental Committee on the Redevelopment of Contaminated Land was formed. In the 1980s, the government took more concerted steps to address the problem, and full regulations were developed in 1995 and implemented through local authorities by 2000.

The regulatory approach in the UK can be described as an integrated one that sees local authorities having responsibility for surveying and registering sites which are thought to be contaminated, while the centralised policy regime encourages integration between technical expertise and the feasible implementation of regulatory principles and concepts across the country (Luo, Catney and Lerner, 2009). The drivers of policy change in this area have included:

- high-population densities and demands for housing
- a need to protect rural heritage, forcing more compact housing
- government identification of areas of growth in brownfield sites
- the establishment of redevelopment in brownfield sites as a key government target.

In this sense, the very act of remediating contaminated land is one which can be considered to be a public service, but delivered by private organisations (developers). The UK’s risk-based management approach to contaminated land has four elements.
1. A national policy framework that is underpinned by an emphasis on risk management and risk-based decision making. It is locally implemented, adopting a ‘suitable for use’ principle and includes remediation as a part of normal development and planning processes. It also provides financial and legislative incentives for remediation and redevelopment, and contains clear guidelines on liability.

2. A regulatory structure which has statutory ‘managers’ at the national level, and primary regulation delivered at the local level by local authorities.

3. Mechanisms through which regulatory capacity is developed and maintained by establishing multidisciplinary qualifications to technical personnel.

4. Provision of financial and legal incentives to developers to remediate certain sites, so that the overall risk is lowered and potential benefits to communities can be delivered.

A few interesting points about the principles underpinning the approach are worth drawing out. First, local authorities are the main implementers of regulation because they know their land best. Second, as mentioned above, the use of financial and legal incentives to encourage development on brownfield land is similar to providing a service or public good for society, and this seems to have enabled positive spillover effects in communities in relation to sustainable development.

There are challenges to the implementation of this system, and interesting governance responses from within the regulated sector have emerged. First, the use of risk-based statutory definitions in the enabling legislation are intended to ensure that national risk management goes hand-in-hand with the local planning process, and that there is integration of the regulatory system. However, some argue that this is not always the case, and misalignments in regulatory philosophies are causing problems and confusion (Catney et al., 2007). For contaminated land, the main regulatory statute within the Environment Act 1995 is problem-solving, project-focused, specific, technical and exclusive, while the planning process is goal-seeking, relational and systems-based. Therefore, there are operational tensions between the different systems which can be problematic and lead to trade-offs being made between political and environmental risk. This situation may be particularly problematic when incentives are used by the regulator to encourage or promote particular areas for redevelopment. If there is local disagreement or technical assessments which end up suggesting that the site should not be developed, tensions can emerge.

Second, the risk management approach rests on technical definitions of contaminants in the soil, called soil guideline values (SVG). These provide an important way of measuring risk. However, the guidelines give a range and arguably are value laden, because they allow flexibility and offer scope for professional judgement to be applied:

It can be argued that SVGs reflect particular societal values, such as how cautious we are as a society and what balance we seek to strike between protecting human health and ensuring continued development. (Luo, 2009, p. 1130)

Third and related to the points above, professionals within the regulated sector have responded to the need to manage different tensions and values by introducing their own certification system. Previously, the sector only required personnel with good qualifications
in the technical fields (science and engineering), but this meant a risk that they would not have the required knowledge (complex and multidisciplinary) to reach informed decisions about the action to take on contaminated sites. In response, professional institutions have come together to set up a professional, specialist qualification (Specialists in Land Condition, SILC). This certification can help reassure all stakeholders that operators are making decisions about safety and planning permission in a way that takes into account diverse sources of information, and expertise which accounts for the views of all stakeholders.

5.3 Environmental governance and proactive corporate environmental management

Environmental risk management today is driven more by a model of environmental governance. This means that the roles of actors are changing. Market actors might assume a greater role in promoting environmental management standards; civic actors might have a more active role in monitoring corporate performance; and governments have to play a more facilitative role for these interactions. In particular, it is argued that in such circumstances, combinations of more traditional command-and-control regulation, together with self-regulation promoted within the industry itself, can be a productive way forward. The appropriate mix can be fine-tuned depending on the particular environmental risks being regulated (Sinclair, 1997).

One example of where this mix has not been appropriately tuned is in the evolution of corporate environmental management and US public policy (Rondinelli and Berry, 2000). It has been asserted that although initial approaches to environmental risk management have benefited from a command-and-control regulatory philosophy to bring environmental risks under control, they were unlikely to have further returns after the initial positive effects. This is because it was likely that corporations were going to see greater returns from adopting pollution prevention initiatives, rather than from simply adopting more ‘end-of-pipe’ measures. They claim that ‘the complex, costly and inflexible command-and-control regulatory system that still dominates environmental policy in the United States neither encourages nor rewards corporate environmental management systems that exceed compliance requirements’ (Rondinelli, 2000, p. 169).

Moreover, the USA suffered from a particular problem of having extremely detailed and complex laws passed by Congress, which required excessive amounts of time to implement, meaning that less time could be spent on identifying how to be more innovative and encourage companies to improve performance. When the Environmental Protection Agency did introduce voluntary programmes designed to incentivise positive performance, it was not very successful. It has been suggested that this was because they were developed by lawmakers, regulators and environmental pressure groups – people with often limited knowledge of business processes and motivations. Therefore, instead of measuring progress in ways that were aligned with business, the objectives of these programmes were designed solely from the environmental perspective, which limited uptake. For example, the Green Lights programme was a partnership with utilities, corporations, non-profit organisations and state, local and city governments to reduce electricity use. The programme measured reduction in carbon dioxide emissions and acres of trees not felled, as opposed to cost
savings to business. If it was framed in terms of the latter, there may have been greater uptake (Rondinelli and Berry, 2000).

Despite the lack of incentivising initiatives on the part of governments, businesses began to adopt their own improved standards and practices. The ISO 14001 environmental management standards were introduced in 1996, with the intention to harmonise environmental management practices across national borders, thereby facilitating international trade. Other benefits of adoption included increased efficiency in production and waste management, reduced risk of potentially costly environmental disasters, lower corporate liability exposure and improved access and competition in the marketplace through more open international trade. By 1999 more than 14,000 corporations had signed up to these standards. However, there was much greater uptake in Western Europe than in the USA (Delmas, 2002).

One reason which has been suggested for this difference in uptake relates to the need for the integration of government action with different governance mechanisms, including those of self-regulation. It is argued that European governments were much more facilitative in supporting the adoption of environmental management standards, by setting up a certification system and providing technical assistance to those wanting to adopt the standards. In the USA, the government did little to encourage adoption, and companies were nervous about joining because adoption meant that they had to make their environmental performance public (Delmas, 2002). Thus, where governments took a more active and facilitative role in supporting companies, there was more appreciation and understanding of their needs and how to encourage positive behaviours. In the long run, this led to enhanced trust between actors.

In a cultural context where there is little trust between ‘polluters’, regulators, environmental non-profit organizations, and investors, it is unlikely that these stakeholders will endorse a standard that does not provide tangible measures of environmental performance. On the contrary, in a context where process is as important as performance, and where the relations between stakeholders are marked by trust, the ISO 14001 standard may find a favorable ground to grow. (Delmas, 2002, p. 98)

In summary, the trends in environmental governance mean that there is an increasing need to adopt a wide range of regulatory mechanisms to encourage positive environmental behaviours. Although corporations have taken their own steps to be proactive in their environmental management, this has been driven by their own needs as organisations to become more efficient, not necessarily out of direct concern for the environment. However, when governments have taken steps to facilitate these behaviours and engage in the process of environmental management from the firms’ perspective, more engagement and better environmental risk management has resulted.

5.4 Environmental governance and stakeholder engagement in the consideration of environmental risk decision-making

The final example of changes brought about by a more dispersed form of environmental governance is found in the move in environmental risk management from technocratic approaches in decision making to more open and deliberative approaches, where stakeholders, corporations and regulators work together (Gouldson, Lidskog and Wester-
Herber, 2007). Of particular interest is that some of this has been initiated by industry, with attempts to engage the public more openly in risk decisions in an attempt to foster confidence and trust, while also determining the best ways to respond to regulatory requirements.

These more inclusive approaches to environmental governance have arisen in response to corporations’ desire to secure more and greater public trust, legitimacy and ‘social licence’ to operate, as well as stakeholders’ and the public’s desire to hold corporations to account. These trends are well established in the literature and have been reviewed comparatively in the context of national and institutional issues which drive these changes. However, what has been less studied – and is more of interest to the comparison in this study – is an analysis of how corporations (the regulated sector) are responding to this new paradigm in environmental risk.

In order to explore these issues empirically, a survey of six organisations involved with a range of complex environmental risk factors was conducted (Gouldson, Lidskog and Wester-Herber, 2007). Gouldson and colleagues sought to understand why companies may or may not adopt more inclusive approaches to environmental risk governance. The study focused on whether ‘opening up built consensus, enhanced legitimacy and led to better management of environmental costs’ or whether it led to greater ‘costs, fuelled conflict and led to compromised decisions’ (Gouldson, Lidskog and Wester-Herber, 2007, p. 60). The organisations were asked questions under three thematic areas: the perceptions that risk managers have of the changing context for risk management; the reasons for and against the adoption of more inclusive approaches to risk management; and early experiences with the adoption of more open and inclusive approaches.

When asked about perceptions of the changing context for risk management, corporate managers reported their thoughts that there were two kinds of risks they had to address: ‘real risks’ and ‘perceived risks’ in society. While they felt that real risks can be addressed through technocratic processes, perceived risks are shaped by social processes and harder to address. However, this does not make them less valid, and in fact there is a healthy body of literature which shows the true value of the lay perspective in expertise.26

Due to these perceptions about the difficulties associated with social risk, there was reluctance in all the organisations surveyed to adopt more inclusive approaches. However, this was overcome in all cases by a perceived lack of legitimacy of the company, thereby forcing companies to seek public engagement. The transition was a difficult process for all companies, which required moving away from more technical approaches to risk management and allowing for more uncertainty to enter the realm of decision making.

When faced with the actual act of engaging stakeholders, risk managers in each company reported various tensions in the actual models or approaches that they needed to adopt, as well as whom or which groups of stakeholders should be engaged. Corporations found it difficult to establish the right mechanisms for engaging with stakeholders in a meaningful way. Where more formal mechanisms were established for engagement, such as town halls or consultations, risk managers found that stakeholders were less likely to engage on the

26 See for example, Wynne (1992) for a seminal article on this topic.
broader issues and instead wanted to push a particular agenda. Conversely, where there was less formal interaction as a result of face-to-face meetings and conversations, more productive interactions were had and the substance behind concerns were aired more readily. However, even when the mechanism was clear, it was not easy to determine the type of information which was supplied to stakeholders. Complex, technical information could be easily misunderstood and so was not always offered; but in any case, stakeholders often were interested in different types of information that risk managers had not always thought of, such as the social or cultural implications of different decisions.

Finally, when asked what their actual experiences were with more open approaches, it was found that most companies initially took an approach of simple information dissemination as a way of communicating, but not necessarily engaging, with the public. Instead of having two-way dialogues, most interactions involved making data more available, publishing reports, using the national media and so on in order to make the public more aware of activities. However, as companies began to develop more meaningful relationships with stakeholders over time, there was generally an improvement seen in both trust and the quality of the relationships between corporations and stakeholders, although this was context-specific. Importantly, where this trust was sustained, social capital built up, and in the long run corporations found it easier to proceed with their business.

5.5 **Lessons from environmental governance**

This chapter has summarised different trends in environmental governance. One of the core issues underpinning environmental governance and driving changes in the sector is the tension between the economic growth of companies and the environmental and health risks to the public. It is the role of the regulator to protect the latter, but it is also in the wider interest of governments to encourage and stimulate economic growth. Several lessons can be drawn from the different examples presented here.

First, it is the case in each example that over time, the role of public trust and confidence has become an important factor in the regulatory system, and that there is merit in exploring the feedback between regulation and governance in this respect. In particular, growing demand from the public for environmental accountability was not only an early driver of regulatory action, but is also a current driver of proactive corporate environmental management. Corporations can be driven to adopt improved environmental standards because of an interest in being seen as having good corporate social responsibility, as well as a need to engage the public more in their decisions about environmental practices in local communities. Although it was initially perceived as a barrier to business processes, proactive engagement with the public led in some cases to improved and lasting relationships that eased tensions, and may have impacted upon corporations’ ability to operate efficiently. In addition, it is not only corporations which have begun to engage the public but also governments, as the notion of expertise has shifted from one that is definitively technocratic to one which is much more broadly based and grounded in lay and local knowledge.
In addition, the public can exercise its voice in other, less direct ways: for example, through not purchasing products from companies that they perceive to have poor environmental standards. There is evidence that labelling schemes have a negative impact on a company’s bottom line, and so the public can be a driver of governance changes in this respect. Harnessing this consumer demand can be a potentially powerful tool in a more dispersed governance system, where the government has less direct control over the behaviour of the regulated sector.

Second, misalignment of regulatory and actor philosophies can pose a threat to the effective implementation of regulation, and hence present challenges for effective and efficient governance responses. In the case of contaminated land use in the UK, there is an active debate about whether this misalignment between local and national regulators does more harm than good. In addition, in the USA the Environmental Protection Agency’s (EPA) experience of introducing voluntary compliance schemes proved ineffective, because they were not designed in a way that took into account the motivations of the business sector, as opposed to those of government or interested stakeholder groups. Understanding the motivations of different actors within a governance system, and working with those motivations rather than not against them, seems to be an important lesson from this sector which can be carried over into other areas.

Related to this is a need to be clear about what kinds of behaviours will be rewarded and by whom. A corporation’s capacity to respond to regulation depends on the capacity to understand the implications of non-compliance and management of related risks. If instead a more collaborative process were encouraged, such as that seen with the diffusion of environmental management standards in Europe, we might move towards a situation where ‘environmental risk may be reconceptualised as the common concern of a group of companies, rather than as a social problem to be ignored and externalized’ (Bennett, 1999, p. 191).

This concept of learning leads us to consider the role of education, training and capacity building as another component of environmental governance. At multiple levels different actors need to engage with each other and the views of different stakeholders in order to create a system that is seen as legitimate and effective. This was the case in contaminated land management, where the industry itself developed a certification scheme for surveyors so that they could be deemed trustworthy and capable of taking multiple viewpoints into account.

Finally, the use of incentives has been shown to be a necessary component of environmental governance, as actors become more dispersed and behaviours more difficult to control. Effective incentives need to take into account the motivations of different actors and be shaped accordingly. Market-based incentives may be appropriate in some situations, but increasingly, more corporate social responsibility-based incentives can be beneficial, as the public finds it has a greater role to play in shaping the system.
CHAPTER 6  Comparative framework 3: financial services sector

Chapter summary

This chapter summarises the regulatory responses to selected key events in the financial services industry in the UK and the USA over the postwar period. Specifically, it considers the Savings and Loan crisis in the USA, the formation and dissolution of the Financial Services Authority in the UK and the global financial crisis, and examines how immediate changes in the regulatory stance affected the behaviour of the regulated sector. A careful analysis of these events suggests that lessons learned included the following.

- Focus less on the form of the regulator and more on consistent application of the rules, along with closure of regulatory gaps. Much has been written about the pros and cons of functional regulation in the USA versus the umbrella regulator in the UK, but both have failed to treat the contributing factors to the global crisis.
- Pay careful attention to potential asymmetry in regulatory effects. Increasing the burden on one segment versus another (such as banks relative to non-banks, in the US case) or one location versus another might lead to inter-segment competition and makes the entire system more fragile.
- Pace changes to regulation and allow adequate time for adjustment. The true impact of changes in regulation are observed only fully with a lag, and this effect is greater, the more complex the regulation. Comprehensive changes should be given longer to mature than simple changes.
- Recognise the trade-off between targeted and comprehensive regulations. Complex regulations (for example, the Financial Services and Markets Act 2000 in the UK and Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 in the USA) may produce the required shifts in firm compliance and risk behaviour, but also may result in overlapping and potentially contradictory rules, making it easier for firms to undertake risky behaviour.
- Carefully consider whether contradictory policy objectives can be executed by a single body, or whether balancing among the objectives should be carried out at a higher level.

6.1 Introduction

At a first level of analysis, financial sector regulation and health sector regulation do not appear to be obvious comparators, as the incentives of key players and the objectives of the enterprises are very different. However, there is some common ground in that both sectors are heavily regulated; in both cases, policymakers are concerned with the regulation of
products offered to the public, and seek to accomplish this via multiple regulatory bodies that oversee a multiplicity of often-overlapping regulations, all of which have implications for operational speed and efficiency in the sectors. Thus, even though the basis of regulation is quite different across the sectors, as typically regulation in healthcare is justified on public goods grounds, while the financial sector is regulated on moral hazard grounds, the authors believe that the financial sector offers a useful comparator.

As in the other comparative frameworks, this study takes an event-led approach to the analysis as the responses of the sector are sharpest after a particular event, which allows regulation-driven changes in firm behaviour to be disentangled from overall trend-level activity.

Financial services firms now operate in a global context. As a result, the sectoral risks are similar across countries. However, national approaches to regulation, rooted in historical and cultural events, are varied. Therefore, the authors have a view on not just specific regulatory actions, but also the context within which those actions appear, which enhances this analysis of the regulated sector’s response. This chapter looks specifically at the UK and the USA, as these provide interesting contrasts in regulatory approach. In both contexts, and indeed in the wider global economy, the trend is toward increasing comprehensive regulation in the industry.

### 6.2 Recent regulatory events in the US financial services sector

The system of financial regulation in the USA remained largely unchanged from the New Deal measures of the 1930s through to the 1980s (Kushmeider, 2005). Regulation is organised along functional lines, meaning that entities are regulated according to the type of product that they offer and is dual-track, meaning that entities face both state and federal regulation in all cases except insurance companies (which face only state regulation), and Government Sponsored Enterprises (which face only federal regulation). The functional emphasis of the system is a direct result of its early beginnings, as in the immediate prewar and postwar periods, financial products were very well defined with little overlap. However, this was no longer the case as the financial system matured.

---

27 Healthcare products are public goods in the economic sense of the term, as these goods and services have large positive externalities (consumption by one individual enhances the well-being of other agents in society). Economic theory suggests that when such externalities are present, private markets typically produce too little of the good or service at too high a price – thus regulation is required to ensure that supply is adequate and access is equitable.

28 ‘Moral hazard’ refers to the phenomenon whereby an insured or otherwise indemnified party takes risks that it would not take in the absence of such protection. Those entities engaged in deposit-taking and loan-making are essential to the functioning of monetary policy. Given this, the state cannot refuse to support them (or their parent companies) in times of trouble. Therefore, those institutions are effectively protected from the potential effects of risky loans and/or investments, and may respond to this protection by increasing risk-taking.

29 Government Sponsored Enterprises are financial services corporations created by the US Congress. They include for housing, the Federal National Mortgage Association (‘Fannie Mae’), Federal Home Loan Mortgage Corporation (‘Freddie Mac’) and Federal Home Loan Banks; for agricultural financing, Federal Farm Credit Banks and Federal Agricultural Mortgage Corporation; and for veteran financial affairs, the Veterans’ Corporation.
The two decades following the Second World War saw significant economic growth in the industrialised world. Along with the associated increase in international trade, this incentivised financial firms to aim for national scale and international consolidation. As financial services entities grew, regulation was relaxed and the incidence of cross-category offerings increased: for example, by the early 1980s a customer seeking a personal loan could approach a bank, Savings and Loan30 or even an insurance company. Yet these entities were subject to a different functional regulator, each of which had its own approach and degree of stringency.

These developments had implications for regulatory effectiveness, although perhaps the gravity of these implications was not fully appreciated until the system was tested by crisis.

6.2.1 Specific event: the Savings and Loan crisis

The 1970s brought significant disruptions in the oil market and, in an attempt to control the consequent inflation, the US central bank, the Federal Reserve, responded by sharply increasing interest rates. The immediate effect was a reduction in net interest margins31 and broader firm profitability. The effect was greatest for Savings and Loans, which was smaller and more heavily focused on loans than other types of financial firms. This tough climate persisted throughout the low-growth, ‘lost’ decade, and by the late 1980s a significant number of thrifts had become insolvent. Indeed, approximately 750 failed thrifts, equivalent to more than 25 percent of total thrifts, were taken over by the government via the Resolution Trust Corporation (Mishkin, 2009). The estimated US$150 billion cost of bailouts associated with this crisis prompted a huge public outcry. In response, the US Congress passed new legislation, the Financial Institutions Reform Recovery and Enforcement Act of 1989 (FIRREA), which abolished pre-existing thrift regulatory agencies, and instead established the Office of Thrift Supervision with increased enforcement powers. FIRREA also established the Federal Housing Finance Board to oversee the federal housing bodies, and increased the enforcement powers of bank regulators. Further, the Act amended the Community Reinvestment Act, which was intended to increase credit to lower income and minority communities. The post-FIRREA US regulatory system is illustrated in Figure 6-1.

---

30 Savings and Loan institutions, also referred to by the broader term ‘thrifts’, are institutions established primarily to encourage personal saving and to facilitate home loans and mortgages. They have their roots in British building societies, as translated by settlers moving to the USA in the 18th century.

31 The net interest margin is the difference between the interest that a financial firm charges on a loan, and the interest that it pays to the depositors or creditors which provide the funds – thus it is a measure of loan profitability.
In the immediate aftermath of the Savings and Loan crisis, thrifts and banks faced more strict regulation than much of the rest of the sector, even though they faced high levels of cross-segment competition. Furthermore, this enhanced regulation was enforced more strictly, as the crisis prompted a heightened awareness of risk and closer scrutiny of the sector among regulators. One example of the firmer application of regulation is that during the decade following the Savings and Loan crisis, the US General Accounting Office issued 13 reports about thrift and banking regulation (Schwartz and Sulitzeanu-Kenan, 2004).

The initial effect was as desired: there was a decrease in risky lending, as both firms and regulators confronted the effects of the crisis. However, by the early 1990s the economy had returned to growth, lifting the performance of the sector. This recovery led to both a return to risky lending and a cautious relaxing of regulators’ post-crisis vigilance. The former was especially true in the non-bank segment, which was able to act strategically to benefit from its post-FIRREA regulatory advantage and capture bank clientele. The ongoing consolidation in the industry led to larger organisations which became even more risk-tolerant as a result of the perceived safety associated with larger balance sheets (the belief that some institutions were simply ‘too big to fail’).

6.2.2 Specific event: the 2007 global financial crisis
The global economy enjoyed a further two decades of sustained high growth from the mid-1990s. Companies began to search out new markets as well as new market segments, and growth in the sector spurred financial innovation and buoyed the wider economy. Policymakers were reluctant to impede this growth, and thus even though restrictions on scope of activity remained formally in effect, regulations tolerated creeping expansion of activity. As a result, this period was one of unprecedented financial conglomeration and cross-segment penetration in financial services. Banking corporations increasingly took on the risk profiles of securities companies as they engaged in trading activity, while securities companies began to take on bank risk as they purchased bank-issued, asset-backed securities. As time passed, the various segments of the industry became increasingly linked

---

32 The Banking Act of 1933, also known as the Glass-Stegall Act, prohibited banks from engaging in securities activities. These restrictions were officially in force until passage of the Graham-Leach-Bliley Act of 1999.

33 As noted in Schooner and Taylor, regulators engaged in ‘liberal interpretations of… Section 20 of Glass-Stegall’ (2003, p. 320).
by this complex web of securitisation, while the boom stimulated consumer and business confidence, leading to more borrowing. Year after year of significant growth also fuelled a relaxation of lending guidelines.

The boom came to an abrupt end with the collapse in the US housing market in early 2007. This triggered a series of foreclosures and failures that devastated the financial system and the real economy over the next several years. Indeed, in the USA alone it is estimated that approximately 460 commercial banks failed during the crisis. More than US$3 trillion has been disbursed by the US government to date in an attempt to shore up the sector. The sheer scale of these bailouts has produced an enormous public backlash, which grew more strident as the crisis triggered a sharp contraction in the real economy.

In response, the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank), the most significant reform of the financial services sector in the USA since the New Deal, was signed into law in July 2010. The Act, which is more than 1,000 pages long, establishes 553 rules and mandates 60 studies as well as more than 90 congressional reports to determine the actual regulatory procedures required to satisfy certain provisions. It increases the powers of the Federal Reserve, US Treasury, Securities and Exchange Commission, Federal Deposit Insurance Corporation and Commodity Futures Trading Commission, and orders the integration of the Office of Thrift Supervision into the Office of the Comptroller of the Currency. It provides for enhanced regulation of large (assets in excess of US$50 billion) and/or ‘designated’ (systemically important) bank holding companies and non-bank financial companies, which for the first time brings key insurance holding companies into the federal regulatory system. Most controversially, the Act provides the Federal Reserve with the authority to restrict the scope of activity of designated financial institutions as a means of reducing systemic risk. Dodd-Frank also establishes the Bureau of Consumer Financial Protection and the Financial Stability Oversight Council, a body composed of representatives of all the major regulators, which will identify systemic risks. The post-Dodd-Frank regulatory system is shown in Figure 6-2.

34 While it is generally accepted that the housing market crash was a leading cause of the crisis, there were many contributing factors. A useful summary is available in Jickling (2009).
35 Federal Deposit Insurance Corporation failed bank list, accessed 6 October 2012.
36 The US government reportedly stands ready to commit up to US$7.7 trillion (Pittman and Ivry, 2008).
The Act has not changed the underlying approach to regulation, which remains functional, although there is more emphasis on systemic risk. Neither has it simplified the regulatory structure: in fact, arguably, by layering the cross-functional Financial Stability Oversight Council and Bureau of Consumer Financial Protection over the existing functional framework, Dodd-Frank has increased the complexity of the system. This addition of an overarching regulatory body to a functioning regulatory structure provides an interesting parallel with the recent restructuring of the UK health governance system.

Many of the reforms called for in the Act will not come into effect for several years, and the extent to which the legislation may leave gaps of interpretation is not yet clear. However, the size and complexity of the proposed changes and the uncertainty surrounding the execution of the required studies have impacted on financial services firms already. Indeed, sectors which have not been relatively seriously impacted by the overhaul currently are positioning themselves to exploit uneven regulation under the Act, and industry representatives have begun lobbying for the most favourable interpretation of legislative language.

In the aftermath of the crisis, financial firms are displaying significantly reduced risk tolerance. In some cases this is a result of the intervention of state control pursuant to bailouts, but in many others this reduction is a consequence of the decline in the risk profile of clients resulting from the deteriorating economic climate. In some sense, then, the reduction in risk-taking in the sector is at least partly a consequence of crisis (so that it would have taken effect without major regulatory action), and of uncertainty resulting from changes in the regulatory environment.

6.3 Recent regulatory events in the UK financial services sector

Historically, the financial services sector in the UK was characterised by a significant degree of self-regulation, with the Bank of England relying heavily on moral suasion rather than explicit regulation. This began to change as the oil shocks and consequent economic
downturn of the early 1970s resulted in several bank failures, and as the European Economic Community moved towards more formal regulation. The first explicit regulation was captured in the Banking Act 1979, which was superseded by the Financial Services Act 1986. The system was formalised further as European directives required. Throughout, the Bank of England retained the triple responsibilities of conducting monetary policy, regulating the financial system and acting as lender of last resort.

As the global economy returned to growth in the early 1990s, financial services firms in the UK, like those in the USA, began to expand areas of operation, thus blurring the lines between product segments. The incoming Labour Government announced in 1997 that it would react to these changes by introducing a single, independent regulator for banking and financial markets, leaving the newly-independent Bank of England to monitor the stability of the financial system and conduct monetary policy. This was the most critical step in breaking with ‘ancient City traditions in favour of a comprehensive and prescriptive statutory-based regulatory regime’ (Norton, 2005, p. 25), and is discussed in more detail below.

### 6.3.1 Specific event: establishment of the Financial Services Authority

The Financial Services Authority (FSA) was established in 1997 as the reconstitution of the Securities and Investment Board, which itself had been established only a decade earlier pursuant to the Financial Services Act 1986. The new entity would operate independently of government but ultimately would report to the Treasury. The Treasury approved the Financial Services and Markets Bill in 1998, but the Act only received royal assent two years later (Norton, 2005). The Financial Services and Markets Act 2000 completely reorganised the UK financial system, positioning the FSA as a nearly-universal regulator, responsible for all financial regulation except those over occupational pension schemes (which is still administered by the Occupational Pensions Authority). The establishment of the FSA gave rise to what was termed the ‘Tripartite Arrangement’ for financial regulation, represented in Figure 6-3.

![Figure 6-3: The tripartite arrangement for UK financial regulation](image-url)
This new structure was intended to increase regulatory effectiveness by allowing the single regulator to have a comprehensive view of system risks and, as mentioned previously, by allowing the Bank of England to focus on conducting monetary policy and maintaining system stability. However, as time went on, four important issues became apparent. First, the FSA introduced a significant number of rules and guidelines within a short period of time. This was not only challenging for the sector, which found it hard to digest these changes and began to view the FSA as intrusive and overly bureaucratic (Norton, 2005), but also for the FSA, since its employees were under pressure to build competence quickly in a raft of new areas. Second, the new regime separated the body to which the firms reported (thus, the body which should have first sight of any institutional challenges) from the body responsible for providing financial stability and support (the Bank of England being the lender of last resort). This opened the way for possible slippages. Third, within the FSA, departments were still organised along functional lines, so that the issue of translating functional view into a system-wide view was only moderated, not completely removed. Finally, the Financial Services and Markets Act 2000 laid out a number of competing objectives for the FSA, including ensuring prudential conduct, promoting competition in the sector and protecting consumers. The FSA was required to balance all of these, elements of which are contradictory.37

The challenges of implementing the vision for the FSA were playing out in the context of a global boom, where the financial sector grew into one of the key drivers of the UK economy. The FSA came under increasing pressure to minimise the burden on UK firms so as not to put them at a competitive disadvantage relative to other countries’ sectors. However, this call for a ‘light touch’ was occurring just as firms, buoyed by optimism from the long boom, were growing more risk-tolerant (Lo, 2008). In this regard, relaxation of regulatory vigilance opened up the way for firms to build holdings of risky assets – a phenomenon which exacerbated the effects of the 2007 financial crisis.

6.3.2 Specific event: The 2007 global financial crisis

The failures in the US property and financial markets were transmitted quickly through the global financial structure, causing serious disruptions in the UK. The government was forced to nationalise the failed Northern Rock and Bradford and Bingley to conduct the orderly wind-up of all obligations, to recapitalise the Lloyds Banking Group and the Royal Bank of Scotland, to lend money directly to functionally insolvent entities, including the Dunfermline Building Society and three Icelandic banks, and to increase the funds available to deposit insurance and consumer protection schemes (National Audit Office, 2010). To date, the UK Government has spent more than £120 billion supporting the banks, and currently has a total contingent liability of more than £510 billion to the sector. The authorities have responded to the weaknesses exposed by the crisis in two ways: by

37 Various elements of the objectives carry direct contradictions. For example, prudential conduct requires managing risk by client, which might call for a financial institution to apply different rules to different consumers; this might fall foul of the non-discrimination aspects of consumer protection. Or indeed, risk management might suggest an optimal size for financial institutions (to carry sufficient overhead at a competitive level), which might lead financial firms to engage in mergers and acquisitions, reducing the number of competitors.
seeking to improve the regulator, and by imposing further restrictions on regulated activities.

On the first point, the government is seeking to increase regulatory efficiency by abolishing the FSA in its current form and sharing its responsibilities between two new entities. The Bank of England will retake responsibility for regulating the safety and soundness of deposit-takers, insurance companies and ‘significant’ investment companies (thereby minimising adverse impacts on financial stability) via the Prudential Regulation Authority, which will be established as a subsidiary of the Bank. The other post-FSA entity will be the Financial Conduct Authority, which will oversee firm conduct in financial markets, trading systems and the regulation of those financial entities not regulated by the Prudential Regulation Authority. In addition, it will take responsibility for protecting consumers. The new bodies are to be established by the end of the current calendar year, and the FSA has reorganised the relevant activities already in preparation for the partition. The new policy is intended to increase regulatory efficiency by allowing the Bank of England to have supervision of the key sectors contributing to systemic risk. The post-FSA restructuring regime is illustrated in Figure 6-4. However, in separating regulatory roles, the new approach risks reverting to the challenges of regulatory overlap, resulting in delays and higher costs to firms. Indeed, there are already reports that authorisations to start new firms are taking on average 14 percent longer to be approved since the FSA has self-partitioned (Masters, 2012).

On the second point, policymakers have reacted to the crisis by revisiting the scope of approved activities for financial firms. In an attempt to ring-fence core financial services (defined as those critical to individual and small and medium-sized enterprise activities), the government may exclude entities carrying out such services from engaging in certain risky investment or lending activities.
6.4 Lessons from financial services

There are several key lessons to be drawn from the events presented above. First, it is clear from the foregoing discussions that even though the regulatory approaches in the USA and UK were very different, the difficulties experienced during the crisis were very similar. This is due largely to the fact that the underlying firm activity (particularly in terms of cross-segment activity) was very much the same, so that the risks were similar. Another contributing factor is that the failures of the divergent systems leave gaps for firms to undertake risky behaviour. In the USA, the main issue seemed to have been overlapping regulators, which left gaps in terms of supervision and accountability (each regulator may assume that the other will flag anything risky); while in the UK, the main issue seemed to have been the competing priorities inside of the single regulator, and the focus on light touch regulation to drive activity in the City. Thus, it would seem that the form that the regulator takes is not as important as the strictness with which regulation is applied.

Second, a further issue is the extent to which asymmetry in regulatory effects within the industry may contribute to further crises. For example, in the USA, FIRREA increased regulation of banks relative to non-bank activities, thereby providing incentives for non-banks to engage in bank-like activities (to capture bank business), as well as for banks to engage in non-bank-like activities (to avoid the costs associated with the regulation). In this way, FIRREA may have contributed to the cross-segment linkages that subsequently made the effects of the 2007 global financial crisis so devastating. Rather than the intended effect of reducing the risks of the system, FIRREA may have had the opposite effect of making the system more fragile.
Third, another lesson relates to the pace of change of regulation and time to adjust in the industry. As mentioned previously, the true impact of changes in regulation are only fully observed with a lag, as the regulated firms take time to understand the impact of the regulatory changes and to craft strategies to minimise their effects. The lesson here is that where possible, changes should be phased so that their true effects can be understood, and policies fine-tuned. It is reasonable to conclude that this effect will be greater, the more complex the regulation, so comprehensive changes should be assessed over longer timeframes than simple changes.

This point links to the degree of complexity of regulation. It is clear that where the regulation is highly complex, as the Financial Services and Markets Act 2000 and Dodd-Frank both were, they require a massive shift in firm compliance and risk behaviour. They also may result in overlapping and potentially contradictory rules which, counter-intuitively, might make it easier for firms to undertake risky behaviour. Moreover, increasing the number of regulators without very clear guidelines about which agency is responsible for which aspects of the policy, may increase the fragility of the system.

A final lesson is that the regulatory approach needs to carefully consider potential contradictions in policy objectives. Where regulation is carried out by a single body, these contradictions must be weighed and objectives prioritised. However, where the contradictions are very large or have significant effects – as is the case with prudential regulation and consumer protection in financial services – perhaps it is not ideal to execute via a single regulator.
CHAPTER 7  Comparative analysis: lessons for research governance in the UK

Chapter summary

This chapter presents the findings of the cross-cutting comparative analysis across all the sectors presented in previous chapters. This analysis focused on three key questions:

Are there common elements or initiatives across the sectors which have impacted upon the way the regulation or regulatory system is received and responded to by stakeholders?

How has the practice of those subject to regulation been affected by those elements, and in what ways?

Can any lessons be extrapolated for the research governance in the UK?

Five common elements or initiatives were identified across the sectors and could be considered further in relation to health research governance in the UK.

1. Increased provision of educational initiatives to improve awareness and training among actors.

2. Transparency and promotion of a culture of openness between researchers, sponsors, trusts, institutions and the public as to the decision making and approval processes which are followed.

3. Together with education and transparency, development of additional mechanisms to foster trust within the system. Formal (accreditation schemes or memoranda of understanding) and informal (networking, relationship-building) mechanisms should be considered.

4. Consideration of the regulatory structure, and where trade-offs may need to be made to align regulatory philosophies and objectives.

5. The use of incentives, in particular the role of the public in creating a demand for research, should be explored. This includes determination of different indicators and metrics against which actors can be evaluated, such as research publications, trials hosted, number of new participants recruited, etc.

7.1 Educational initiatives

Education, transparency and trust are all closely linked, as the two former elements lead to trust. However, they were sufficiently distinct in the literature reviewed across sectors to merit being drawn out separately before being brought together. The importance of education and the presence of educational initiatives which benefit all actors in a regulatory system was a common theme across all sectors, although they played slightly different roles
in each one. In the analysis of health research systems, we saw that education could play an important role in helping stakeholders to distinguish between the actual and perceived risks of allowing for more centralised review systems. This was particularly notable in Australia, where case studies found that people were basing reviews on the theoretical risk of other review boards not doing as complete an assessment, while evidence has shown that more review does not lead either to more ethical research, or to greater patient safety (Beagan and McDonald, 2005). On a related point, there was evidence in Australia, Canada and the USA of a lack of robust, empirical evidence about the outcomes of research review processes; in particular, a lack of commonly agreed outcome metrics about what a ‘good’ review process meant, and how it could be measured.

In the environmental sector, educational initiatives, which involved the provision of information and support for regulatory implementation, were shown to be an effective way to enhance the way that corporations took up new initiatives, such as environmental performance standards. The introduction of product-labelling schemes has been effective in promoting consumer-driven demand, which in turn has driven more positive environmental behaviours on the part of industry.

Education and information provision have been crucial to successful implementation in the financial sector, in light of new regulatory frameworks introduced after the 2007 financial crisis. Unsurprisingly, the more complex the regulation, the more important it is for each party to understand the scope of the activities for which they are responsible.

In all of these examples we see that the purpose of education is often to improve the transparency of a process and to increase trust and understanding among actors. The contribution of different educational initiatives for research governance is to foster joint learning and understanding among actors (including NHS Trusts, review boards, researchers, sponsors and the public) about their different motivations and expectations.

7.2 **Transparency among actors**

Like education, transparency is another way of achieving trust. Although this did not emerge as a key theme in all the sectors, it was prominent in those where it did. In the different health research systems, an overall culture of openness will encourage trust, as people feel that they are part of the bigger process and that the whole process is accessible to them. It has been observed that when committees were able to see the full spectrum of information that others had used to reach a decision, this helped in trusting others’ conclusions. Initiatives such as IRBShare in the USA are trying to promote transparency in the review process through the open sharing of approvals and documentation throughout the process. Another way of promoting transparency in decision-making processes is to adopt standard operating procedures (SOPs), which can be used as a measure to reduce the uncertainty in others’ decisions. Since a common set of criteria are applied to all reviews, everyone is certain about the process used, and this can create confidence and trust.

More inclusive and open approaches to deliberating on different environmental regulations and approaches were viewed as a way of including a variety of actors, and therefore promoting transparency in decision making. This can lead to more positive interactions with the system on both sides. As regards the environment regulation literature reviewed,
there were examples of this behaviour coming from corporations, and in this case the
government can play a facilitative role in enabling this behaviour.

7.3 **Trust and confidence in the system and system actors**

Both education and transparency contribute to the building of trust and confidence between actors. Trust emerged from the present analysis as a common element and important feature across regulatory systems and cultures which could lead to improved responsiveness and uptake of different types of regulation. For example, in the financial sector the relationship between the regulator and the sector was key, and when it was adversarial, such as is the case after the 2007 crisis, then the process for implementing regulation and achieving effective and efficient compliance is longer. Trust can be built up between different actors within the wider governance system, not just between the regulator and the sector. Public trust emerged as an important driver of behaviour in the case of environmental regulation, as it was the desire to obtain public trust that drove some companies to engage proactively in environmental management. In the medical sector, public trust drove the regulator to be more accountable and hence provide more consistent guidelines. In addition, trust can be seen as a prerequisite for a regulatory framework to operate effectively. For example, for research institutions involved in multi-site research, trust between institutions is required so that review boards are able to trust the decisions made by, and accept approvals from, other review boards.38

Different ways of achieving trust which were observed in the literature, including educational initiatives and greater transparency, as discussed above. Another mechanism related to education is through accreditation. Australia uses an accreditation system as a formal way of instilling trust in review boards which receive decisions from other review boards; and in Canada, official agreements are used between institutions to formalise and effectively ‘declare’ a level of trust in the process and acceptance of another’s review. In the US health research system, national-level training is offered by an accreditation organisation; and in the environmental sector, certification was used as a way of instilling trust in contaminated land inspectors.

Furthermore, clear application of the regulations themselves is important. In the medical drug sector, this study found that regulators need to develop clear, consistent rules and appropriate thresholds for the approval process, in order to build the trust of product sponsors as well as the general public. In addition, regulators needed clear, consistent strategies for dealing with ‘failures’, so that individual poor outcomes do not erode trust in the whole system.

Less formal mechanisms can be used to foster trust. In the USA, the National Cancer Initiative is exploring ways of working with institutions on an informal basis to encourage their acceptance of centralised review board decisions. Increasing interaction and communication between the committees would remove the facelessness of another institution and, through these interactions, could create trust.

---

38 Ethics review as a component of institutional approval for a multicentre continuous quality improvement project: the investigator’s perspective.
7.4 Structure

Different regulatory structures were observed in all systems, and there were implications for actors’ responsiveness in the system as a result of these differing structures. However, these structural issues varied in nature, and do not imply that a wholesale change in regulation is required to address them. They could be as simple as the timing of the review process in different health systems (for example, sequential versus parallel reviews), or they could be related to the use of targeted versus more comprehensive regulatory frameworks.

In Australia the various stages of the review process occur sequentially, ensuring that others are already in possession of previously completed approvals before commencing their own reviews. Duplication could be reduced through the introduction of such step-wise processes, so that each party could be confident that certain decisions which were the responsibility of others had been made already, and did not need to be looked at again.

The nature of regulation is a more fundamental issue, and not always something that can be easily changed; nevertheless, it is important to consider for its implications. While targeted regulations are quicker and easier to understand, they are only effective if a few causal factors of a problem can be readily identified. Comprehensive regulation can produce wholesale changes in behaviour and eventually culture, but takes time to be digested by the sector. If not carefully set out, it may leave gaps of interpretation which can be exploited. This was seen in the financial sector, and similarly in environmental regulation. In the latter case, more top-down, source-based regulations were quick to be implemented, but as the complexity of the system increased, a more dispersed governance model of regulation was required, which relied much more on incentivising positive behaviour and creating comprehensive enabling environments for the sector to adopt regulation. Moreover, the analysis of the financial sector concluded that regulation must be written to reflect the changing reality of the marketplace, suggesting that an awareness of the motivations and environments that the sector operates within is critical.

In addition to the nature of regulation, a clear and common theme was the need for a common set of regulatory objectives to be clearly communicated and aligned to different actors in the system. Where objectives are in conflict, problems can occur and attention should be paid to how to overcome these. For example, in the medical sector, the objectives of the FDA in the USA and the National Institute for Health Research and Clinical Excellence (NICE) in the UK are to provide quick and/or high availability of medicines (and at low cost in the case of NICE), and to protect public safety. Sometimes these two aims can be in conflict, and the regulator must protect itself against erring too far to one side. In the financial sector, too, internally inconsistent objectives can lead to a situation where policy goals are in direct conflict (such as where an emphasis on prudential guidelines may lead firms to merge and thereby reduce competition). If this is the case, trade-offs must be set out at the appropriate level and regulatory operation must be structured so as to minimise these conflicts. In the environmental sector, misalignment of regulatory philosophies was seen when national regulations were implemented at a local level. This misalignment can be a real barrier to effective regulation.
7.5 **Incentives**

All the sectors reviewed, except for the research systems, showed successful use of either positively or negatively oriented incentives to affect regulatory behaviour. In the financial sector, such mechanisms have serious effects on industry culture. In the USA, for example, regulation has been formal and compliance mechanisms traditionally have been ‘sticks’ or negative incentives (‘do not comply and you will be punished’), whereas until the late 1970s in the UK, regulation was informal and moral suasion or ‘carrots’ were used. However, ultimately it is the public that are clients and in many cases shareholders, and public perception can have commercial consequences if not addressed. Thus public pressure can be used to a regulator’s advantage as an incentive to encourage compliance.

The public has played an incentivising role in the case of environmental regulation. Here, more traditional incentives have been economic and/or market-based, and are commonly used to promote positive corporate behaviours, usually with the intention of encouraging self-regulation and reducing the cost of more direct regulation by the state. These are most effective when tailored to the needs of business, or those whose behaviour the regulatory framework is trying to affect. In recent years, growing pressure from the public and stakeholder groups to encourage environmental responsibility has led to a more proactive stance on environmental management by companies themselves, including a greater use of corporate social responsibility initiatives and integration of environmental standards. The regulator can facilitate and promote this reciprocal relationship by creating a regulatory environment that further incentivises positive behaviours, or sends clear signals which may increase the adoption of standardisation.

In a similar fashion, medical drug firms are disciplined to a large extent by the ‘court of public opinion’ as they face serious commercial and legal consequences (civil payouts) if a flagship product has adverse effects. Well-organised lobby groups have been shown to change the mechanisms of regulatory action (for example, the AIDS lobby influencing accelerated approval in the USA).

What is important to note from all of these examples is that it is important to have a clear sense of who the ultimate ‘customers’ of the regulation are. In the cases here, the public and other stakeholder groups have played a role in incentivising more reactive and responsive behaviours to regulatory frameworks. In the health research system, ultimately there are many different customers of the regulation, from sponsors and researchers to NHS Trusts and the public. Thinking about how each of these actors could be used to incentivise different types of behaviour might improve research governance. For example, increasing public awareness of the importance of participating in research could encourage them to demand that trials be run out of their local NHS Trusts, so they can take part.

7.6 **Final remarks**

In the health research sector, governments want institutions and, where appropriate, local NHS Trusts to undertake research activities that will lead to improvements in the public’s health and, ideally, improved and more efficient delivery of health services. This in itself can lead to economic benefits as well as health and social ones. However, there are observed barriers to an efficient research governance system, and this work has sought to
explore other health research systems as well as other regulated sectors, to determine what mechanisms can be used to improve the responsiveness and receptiveness of various actors to regulation.

In this study, seven health research systems were reviewed: three in greater detail for their attempts to incorporate multi-centre review into their systems; and three additional regulated sectors – medical drugs, environmental management and the financial sector. It was found that although specific regulatory mechanisms varied within each sector, there were common themes that could be identified and which could be explored further by the DH for their applicability to research governance:

- increased provision of educational initiatives to improve awareness and training among actors
- transparency and promotion of a culture of openness between researchers, sponsors, trusts, institutions and the public as to the decision making and approval processes that are followed
- together with education and transparency, development of additional mechanisms to foster trust within the system – formal (accreditation schemes or memoranda of understanding) and informal (networking, relationship-building) mechanisms should be considered
- consideration of the regulatory structure and where trade-offs may need to be made to align regulatory philosophies and objectives
- the use of incentives, in particular the role of the public in creating a demand for research, should be explored – this includes determination of different indicators and metrics against which actors can be evaluated, such as research publications, trials hosted, number of new participants recruited, etc.


